Tumor-Specific Recognition Molecules

The invention relates to recognition molecules which are directed towards tumors and can be used in the diagnosis and therapy of tumor diseases.

Tumor diseases or cancerous diseases are oncotic diseases which can be described by a locally confined increase of tissue volume. In a broader sense, any localized swelling as a result of oedemas, acute and/or chronic inflammations, an aneurysmatic expansion or even organ swelling caused by inflammation is a tumor. More strictly speaking, especially formation of new tissue such as tumescence, blastomas and/or neoplasias in the form of a spontaneous, variably disinhibited, autonomous and irreversible excessive growth of autologous tissue, normally associated with more or less distinct loss of specific cells and tissue functions, is understood to be a tumor disease. Tumors can be systematized according to their biological behavior, but also into a histogenetic taxonomy, or according to clinical or pathological findings.

Specifically in the clinical sector it may be necessary to recognize tumors as early as possible and in a selective fashion as well, because early recognition and the treatment or removal that follows will ensure successful treatment of the swelling without deformation of the affected organ structures or gene sections, thereby also preventing formation of metastases. In subsequent examinations following a cancer treatment even slightest metastases must also be detected at an early stage in order to optimize further aftercare. In many sectors of occupational medicine and health care it is also necessary to determine whether a tissue or an organ has potential susceptibility to cancer

before the organ or tissue has already undergone degeneration or transformation.

The oldest and - at the same time - simplest method of tumor recognition sometimes used successfully even today is palpation and visual observation. Thus, for example, mammary carcinomas or prostate carcinomas are palpable as nodes. Indications of skin cancer as a result of conspicuous birthmarks can be detected optically by physicians or patients themselves. Other optical procedures are imaging methods, for example, wherein images of the body are recorded by means of apparatus, in which images a tumor can be recognized. These methods include e.g. X-ray irradiation, as well as computer tomography (CT). In these procedures the body is screened with high-energy radiation, and the degenerate tissue structures can be recognized as a result of the transparency change for such radiation compared to healthy tissue. Frequently, contrast media are used in such methods, which are injected into the corresponding regions, increasing the absorption. In addition, cancer diagnosis is possible by means of ultrasound or by using radio-labelled antibodies, in which case the tumor-typical antigens will bind to the organs to be examined, so that the tumors can be recognized in the imaging procedure. In addition to imaging methods, laboratory investigations are another important means of early detection of cancer, where samples of urine, blood or tissue are examined for abnormal features. For example, this might be an altered composition of such samples, but also, appearance of substances normally not occurring or only in small quantities. These substances are generally referred to as tumor markers. They are either produced by the tumor tissue itself or formed as a body response to the tumor. In addition to substances, cellular changes whose qualitative or quantitative analysis allows a statement as to the presence, course or prognosis of malignant diseases are also referred to as tumor mark-

ers. Most tumor markers are physiologically occurring or modified substances which can be detected in urine, serum or other body fluids at higher or lower concentrations compared to physiological conditions or normal genotypical/phenotypical expression, or in or on tumor cells, said substances being synthesized and/or secreted by the tumor tissue and subsequently liberated by tumor decay or formed in response of the organism to a tumor. A wide variety of tumor markers has been described, the use of which is considered reasonable especially in colon cancer, breast cancer, ovary cancer, prostate and testicle cancers and in small-cell lung carcinoma. Such cancer markers include e.g. CEA, CA 15-3, CA 125, α-fetoprotein, HCG, prostate-specific antigen, neuron-specific enolase, CA 19-9 and SCC.

By an increase in serum or in tissues or by their presence as modified proteins, lipids and/or carbohydrates, the above-mentioned markers, on the one hand, indicate e.g. (i) inflammatory diseases, intestinal polyps, viral inflammations and, on the other hand, especially (ii) cirrhoses, degenerations, tumors and metastases. A major part of these markers consists of molecules comprising both protein and carbohydrate structures, and possibly lipids. The lower the protein level and thus, the higher the carbohydrate or lipid level of these markers, the more difficult is detection thereof using e.g. recognition molecules such as antibodies. Up to now, various antibodies to carbohydrate structures have been produced by immunization of mice using the hybridoma technology.

Cancer diagnostics using recognition molecules involves several disadvantages. Thus, certain tumor markers may also be present in non-cancerogenic diseases, so that the recognition molecules employed indicate a positive reaction. Furthermore, non-interaction of recognition molecules does not indicate the absence of a tumor disease. Another draw-

back is that well-known recognition substances are normally non-specific. That is, positive detection rarely indicates a specific type of tumor disease. In addition, another and crucial drawback of well-known recognition molecules is their limited usability in monitoring the development of tumors, e.g. subsequent to surgery. As a rule, the use of well-known tumor markers therefore is not possible in early recognition or in aftercare, especially in prophylaxis.

In addition to the above general disadvantages, there are some specific drawbacks in recognition molecules directed towards carbohydrate structures. Immunization with carbohydrate antigens usually results in a primary IgM response only, or immune response is completely absent because many carbohydrate structures are also autoantigens. Carbohydrates are T cell-independent antigens incapable of inducing class switching and associated maturing by somatic mutations, which is why the antibody response is usually restricted to the IgM class. Therefore, due to the generally weak interaction and necessary multivalence, it is difficult to produce high-affinity antibodies. One problem with antibodies to carbohydrate structures not only is low affinity, but also the specificity. In particular, production of specific antibodies to short uncharged carbohydrate structures is extremely difficult, and in many cases a certain specificity is only achieved when the carbohydrate structure is localized on a specific carrier. Thus, for example, the JAA/F11 antibody which is directed towards Galβ1→3GalNAc not only recognizes this antigen, but also GlcNAcβ1→6Galβ1→3(GlcNAcβ1→6)GalNAc and - although with lower avidity - Galβ1-3GlcNAc. More recent ways of obtaining recognition molecules using various forms of combined techniques, such as phage display technology, neither solve the above-mentioned disadvantages. The problem of weak recognition molecule-carbohydrate interaction remains in this latter case as well. In this context, particular attention

should be given to the fact that the primary IgM antibodies which are the most frequent ones obtained by immunization are too large in size for therapeutic use. Another disadvantage of well-known recognition molecules for tumor markers is that they do not make the tumor recognizable until it has already reached a critical size. That is to say, early stages of tumor growth cannot be determined with well-known recognition molecules directed towards tumor markers.

Another drawback of well-known recognition substances is that "functional" use thereof is not possible. "Functional" means that the recognition molecules bind to the tumor markers not only in such a way that the latter are detected, but that they interact with the tumor cell via markers in such a way that the tumor cell is impaired in its growth. Such recognition molecules may specifically interact with particular tumor markers, which are immobilized e.g. on the surface of tumor cells, in such a way that the tumor characterized by the tumor markers is given a therapeutic treatment. On the one hand, these functionally active recognition molecules are capable of detecting tumor cell-associated tumor markers and, at the same time, prevent the tumor cell from further growth or formation of metastases as a result of binding to this tumor-specific structure. Disadvantageously, well-known recognition molecules are capable of affecting tumor growth only in rare cases. As a rule, additional substances restricting or inhibiting tumor growth therefore must be coupled to the antibody, so that the latter represents the "shuttle" of said substance rather than the agent of treatment.

The object of the invention is therefore to provide recognition molecules which, on the one hand, allow easy, reliable and efficient detection of tumors and, in addition,

can be used in the prophylaxis, therapy and/or aftercare of tumors.

The invention solves the above technical problem by providing recognition molecules comprising an amino acid sequence which contains the amino acid sequence SEQ ID No. 1 and the amino acid sequence SEQ ID No. 2 or 3 and the amino acid sequences SEQ ID No. 4, 5 or 6, said recognition molecules specifically binding the core 1 antigen.

Mutatis mutandis, the definitions of terms given below also apply to statements given above, those given here and hereinafter.

According to the invention, the term recognition molecule is understood to concern a molecule which, especially under stringent conditions, specifically binds the core 1 carbohydrate structure.

According to the invention, core 1 is understood to be the carbohydrate structure Galβ1-3GalNAc which can be present as α-anomer (Galβ1-3GalNAcα) or β-anomer (Galβ1-3GalNAcβ). Preferred in this context is the α-anomeric variant. However, the recognition molecules according to the invention can also bind the α-anomer Galβ1-3GalNAcα alone or both anomers Galβ1-3GalNAcα and Galβ1-3GalNAcβ in the same way.

According to the invention, specific binding towards core 1 is understood to be binding that recognizes core 1 only, preferably the α-anomer, or recognizes core 1 and core 2 (Galβ1-3(GlcNAcβ1-6)GalNAcα). The recognition molecules do not exhibit any cross-reactivity with other derivatives and anomers of carbohydrate structures such as given in Example 7. The recognition molecules of the invention do not interact with Galα1-3GalNAcα, Galα1-3GalNAcβ, GalNAcα, Neu5Acα2-3Galβ1-3GalNAcα, Galβ1-3(Neu5Acα2-6)GalNAcα, GlcNAcβ1-

2Galβ1-3GalNAcα, GlcNAcα1-3Galβ1-3GalNAcα, GalNAcα1-3Galβ and 3'-O-Su-Galβ1-3GalNAcα under the conditions described in Example 7. In particular, determination is effected by means of specificity tests using well-defined synthetic carbohydrate structures.

In a preferred embodiment a recognition molecule of the invention specifically binding the core 1 antigen comprises:

a) a first amino acid sequence which contains the amino acid sequence SEQ ID No. 1 and the amino acid sequence SEQ ID No. 2 or 3 and the amino acid sequence SEQ ID No. 4 or 5 or 6; and

b) a second amino acid sequence which contains the amino acid sequence SEQ ID No. 7 or 8 or 9 and the amino acid sequence SEQ ID No. 10 or 11 and the amino acid sequence SEQ ID No. 12 or 13.

The first and the second amino acid sequence can be present on one or more and preferably two polypeptides.

The core 1-binding recognition molecules according to the invention are characterized in that a defined set of single amino acid sequences is included therein. The amino acid sequence of said recognition molecules includes one or two triplets of defined sequences. These sequences represent the binding domains and define the specificity of the recognition molecules. The 1-triplet recognition molecule contains the amino acid sequence SEQ ID NO. 1, the amino acid sequence SEQ ID NO. 2 or 3 and the amino acid sequence SEQ ID NO. 4 or 5 or 6. Core 1-specific recognition molecules defined by two triplets contain the amino acid sequence SEQ ID NO. 1, the amino acid sequence SEQ ID NO. 2 or 3 and the amino acid sequence SEQ ID NO. 4 or 5 or 6 for the first triplet, and the amino acid sequence SEQ ID NO. 7 or 8 or 9, the amino acid sequence SEQ ID NO. 10 or 11 and the amino acid sequence SEQ ID NO. 12 or 13 for the second

triplet. The first and the second triplet can be present either on one or on more polypeptide chains which, in the latter case, together form the binding recognition molecule. Further, in the meaning of the invention, these triplets are referred to as triplet sequence 1 for the first amino acid sequence being included and as triplet sequence 2 for the second amino acid sequence being included; see definition a) and b) of the description above. According to the invention, the recognition molecule can be an antibody, particularly a murine, chimeric or human IgG or IgM, an scFv structure or other.

Another embodiment of the invention relates to recognition molecules wherein at least one amino acid sequence of SEQ ID Nos. 1 to 13 has been modified by mutation, deletion and/or insertion, but wherein the property of binding specificity towards core 1 continues to exist. Advantageously, this is utilized to improve the recognition molecules, e.g. with respect to affinity, solubility and/or producibility.

In a preferred embodiment, modification of a recognition molecule is effected by one or more mutations in one or more amino acid sequences selected from SEQ ID Nos. 1 to 13, wherein single amino acids are replaced by amino acids having analogous physicochemical properties which, advantageously, do not fundamentally change the three-dimensional structure of the binding domain in the recognition molecules, so that the core 1 specificity of the recognition molecules is retained. Amino acids having analogous physicochemical properties in the meaning of the invention can be summarized into 6 separate groups and are illustrated in Table 1.

Table 1:
Amino acids with analogous physicochemical properties regardless of molecular size

Property or functional group	Amino acid
aliphatic	glycine
	alanine
	valine
	leucine
	isoleucine
hydroxy group	serine
	threonine
carboxyl group	aspartic acid
	glutamic acid
amide group	asparagine
	glutamine
amino group	lysine
	arginine
aromatic	phenylalanine
	tyrosine
	tryptophane

In another preferred embodiment of the recognition molecules of the invention specifically binding core 1, at least one amino acid sequence of amino acid sequences SEQ ID Nos. 1, 2, 3, 7, 8 and/or 9 is replaced by canonical structure variants or equivalent structures having the amino acid sequences SEQ ID Nos. 14 to 45, with SEQ ID NO. 1 being replaced by a sequence of sequences SEQ ID Nos. 14 to 17 (CDRH1), SEQ ID NO. 2 or 3 by a sequence of sequences SEQ ID Nos. 18 to 27 (CDRH2), and SEQ ID NO. 7 or 8 or 9 by a sequence of sequences SEQ ID Nos. 28 to 45 (CDRL1).

The general relationship between an amino acid sequence and
the tertiary structure of loops formed by these sequences
is well-known to those skilled in the art and has been in-
vestigated in detail [Rooman et al., 1989; Martin, Thorn-
ton, 1996]. Immunoglobulins represent a unique example. By
analyzing the loop conformations of the hypervariable re-
gions (complementarity determining regions, CDRs) in the
light and heavy chains of antibody molecules, so-called ca-
nonical classes have been defined [Chothia, Lesk, 1987;
Chothia et al., 1986, 1989, 1992; Wu, Cygler, 1993]. On
this basis, the canonical structure variants SEQ ID Nos. 14
to 45 of SEQ ID Nos. of 1, 2, 3, 7, 8 and 9 have been de-
rived.

The amino acid sequences SEQ ID Nos. 1 to 13 or their modi-
fications in a core 1-specific recognition molecule in the
meaning of the invention form spatial structures, e.g. so-
called loops which are characterized by possessing a defin-
able tertiary structure and/or quaternary structure. The
binding region of a recognition molecule with the core 1
antigen is formed by amino acid residues which are provided
by up to six variable loops on the surface of the molecule
and specifically interact with core 1.

In another embodiment of the invention, recognition mole-
cules specifically binding core 1 are provided, wherein at
least one sequence of the triplet sequences is omitted,
which is not immediately involved in the interaction with
the core 1 antigen.

In another embodiment the recognition molecules comprise at
least one of the amino acid sequences SEQ ID Nos. 1 to 13
or the above-described variants thereof in duplicate or
multiplicity, and such doubles may also be present in the
form of variants of the same amino acid sequence. All rec-
ognition molecules described in this section advantageously

recognize the core 1 antigen in a specific manner. For eas-
ier comprehension, the above recognition molecules as well,
which, strictly speaking, do not bear any triplet sequences
as a result of omitting or multiplying sequences, will nev-
ertheless be referred to as triplet sequence 1 or triplet
sequence 2 hereinafter.

In another embodiment the recognition molecules of the in-
vention specifically binding the core 1 antigen comprise
amino acid sequences having a homology of at least 60%,
preferably 70%, more preferably 80%, especially preferably
90%, with respect to the sequences SEQ ID Nos. 1 to 13.

Furthermore, the recognition molecules in the meaning of
the invention may comprise framework sequences which sepa-
rate the comprising amino acid sequences, i.e. amino acid
sequence SEQ ID NO. 1 and amino acid sequence SEQ ID NO. 2
or 3 and amino acid sequence SEQ ID No. 4 or 5 or 6, or the
above-described variants thereof, and framework sequences
which separate the amino acid sequence SEQ ID No. 7 or 8 or
9 and the amino acid sequence SEQ ID No. 10 or 11 and the
amino acid sequence SEQ ID No. 12 or 13, or the above-
described variants thereof. The first and the second amino
acid sequence can be present on one or more and preferably
two polypeptide chains. In the meaning of the invention,
such framework sequences are also referred to as spacers
and may vary in length and sequence. This expressly in-
cludes those recognition molecules wherein not all of the
amino acid sequences SEQ ID Nos. 1 to 13 or the above-
described variants thereof are separated by spacers. More-
over, the recognition molecules preferably have additional
flanking amino acid sequences likewise referred to as
framework sequences in the meaning of the invention.

More specifically, the framework sequences have the func-
tion of forming the above-described amino acid sequences

responsible for or involved in core 1-specific binding of
the recognition molecules into a suitable configuration and
spatial structure so as to allow binding to core 1. It can
be envisaged that the amino acid sequences SEQ ID NO. 1 to
NO. 13 without at least one additional amino acid sequence
as framework sequence are incapable of binding the core 1
antigen in a specific fashion in the meaning of the inven-
tion. Moreover, the framework sequences may provide the
recognition molecules with e.g. the required biological and
chemical stability, so that the spatial structure can be
built up effectively and maintained for function and use in
a suitable functional form which includes core 1 binding.

In a preferred embodiment the triplet sequences are intro-
duced in existing proteins by replacement of amino acid se-
quences and/or by addition, the existing protein sequences
serving as framework sequences in the meaning of the inven-
tion, or framework sequences being taken from suitable pro-
teins. For example, such framework sequences can be modi-
fied by means of mutations, deletions or insertions. Meth-
ods of molecular biology, biochemistry and protein engi-
neering *per se* known to those skilled in the art can be em-
ployed for this purpose. Preferred proteins for this pur-
pose are proteins of the immunoglobulin superfamily, prote-
ase inhibitors, lectins, helix bundle proteins and lipo-
calins, such as disclosed in: Nygren and Uhlen, 1997; Nut-
tall SD et al., 1999; and Skerra, 2000.

In another preferred embodiment the framework sequences are
antibody framework sequences from one or various species or
amino acid sequences mimicking the consensus sequence of
framework sequences of murine, human antibodies and/or an-
tibodies of other mammals. A consensus sequence is an ide-
alized sequence wherein the most frequently occurring amino
acid is representative in each position when comparing a
large number of existing sequences, e.g. from antibody data

bases. The recognition molecules preferred herein are characterized in that the framework sequences for the first triplet sequence 1 comprising the amino acid sequence SEQ ID NO. 1, the amino acid sequence SEQ ID NO. 2 or 3 and the amino acid sequence SEQ ID NO. 4 or 5 or 6, or the above-described variants, are antibody framework sequences of the variable heavy chain, V_H, in the literature also referred to as framework sequences, and the framework sequences for the triplet sequence 2 comprising the amino acid sequence SEQ ID NO. 7 or 8 or 9, the amino acid sequence SEQ ID NO. 10 or 11 and the amino acid sequence SEQ ID NO. 12 or 13, or the above-described variants thereof, are antibody framework sequences of the variable light chain, V_L.

Also preferred are antibody framework sequences of antibodies from mammals, with antibody framework sequences of human and/or murine origin being particularly preferred. The framework sequences can be combined from antibody framework sequences of various species. Such antibody framework sequences are well-known to those skilled in the art and can be obtained from various data bases such as the Kabat data base (immuno.bme.nwu.edu) or the National Center for Biotechnology Information data base (www.ncbi.nlm.nih.gov). Likewise, these antibody framework structures can be extended by additional amino acids and/or modified by one or more mutations, e.g. deletions and/or insertions, with specific binding to core 1 being retained.

When combining the triplet sequences with antibody framework sequences in a preferred variant of the invention, the recognition molecule represents a variable chain of an antibody or a structure derived therefrom.

Particularly preferred antibody framework sequences as framework sequences in the meaning of the invention are the amino acid sequences corresponding to FRH1, FRH2, FRH3 and

FHR4 in Table 2 for the variable heavy chain and the amino acid sequences corresponding to FRL1, FRL2, FRL3 and FRL4 in Table 2 for the variable light chain, the amino acid sequences of the triplet sequences 1 and 2 with SEQ ID Nos. 1 to 13 corresponding to the corresponding CDR regions of the antibodies. The variable heavy (V_H) and light (V_L) antibody chains, respectively, are composed as follows: V_H: FRH1-CDRH1-FRH2-CDRH2-FRH3-CDRH3-FRH4, and V_L: FRL1-CDRL1-FRL2-CDRL2-FRL3-CDRL3-FRL4. Table 2 illustrates the positions in detail. The positions of the individual amino acids or amino acid sequences correspond to the numbering of amino acids in antibody molecules according to Kabat.

Table 2:

Name	Position range	Pos.	Amino acid or amino acid sequence
FRH1	1 to 30	1	Q or E
		2	V
		3	Q, K or T
		4	L
		5	K or V
		6	E or Q
		7	S
		8	G
		9	A
		10	E
		11	L or V
		12	V or K
		13	R or K
		14	P
		15	G
		16	T or A

		17	S
		18	V
		19	K
		20	I or V
		21	S or P
		22	C
		23	K
		24	A, V, S or T
		25	S
		26	G
		27	Y, F, S or D
		28	T
		29	F, L or I
		30	T
CDRH1	31 to 35		SEQ ID NO. 1 and variants
FRH2	36 to 49	36	W
		37	V
		38	K or R
		39	Q
		40	R or A
		41	P
		42	G
		43	H or Q
		44	G
		45	L
		46	E
		47	W or R
		48	I or M
		49	G

CDRH2	50 to 65, with position 52a introduced in addition		SEQ ID NO. 2 or 3 and variants
FRH3	66 to 94	66	K or R
		67	A or V
		68	T
		69	L or M
		70	T
		71	A, L or T
		72	D
		73	T
		74	S
		75	S or T
		76	S
		77	T
		78	A
		79	Y
		80	M
		81	Q or E
		82	L
		82a	S
		82b	S or R
		82c	L
		83	T or R
		84	S
		85	E
		86	D
		87	S or T
		88	A
		89	V
		90	Y

		91	F or Y
		92	C
		93	A
		94	Y, K or R
CDRH3	95 to 102, with positions 100a and 100b introduced in addition		SEQ ID NO. 4, 5 or 6 and variants
FRH4	103 to 113	103	W
		104	G
		105	Q
		106	G
		107	T
		108	T, S or L
		109	V or L
		110	T
		111	V
		112	S
		113	S or A
FRL1	1 to 23	1	D
		2	I, V or L
		3	Q or L
		4	M
		5	T
		6	Q
		7	T or S
		8	P
		9	L
		10	S
		11	L
		12	P
		13	V

		14	S or T
		15	L or P
		16	G
		17	D or E
		18	Q or P
		19	A
		20	S
		21	I
		22	S
		23	C
CDRL1	22 to 34, with positions 27a, 27b, 27c, 27d and 27e introduced in addition		SEQ ID NO. 7, 8 or 9 and variants
FRL2	35 to 49	35	W
		36	Y
		37	L
		38	Q
		39	K
		40	P
		41	G
		42	Q
		43	S
		44	P
		45	K or Q
		46	L
		47	L
		48	I or V
		49	Y
CDRL2	50 to 56		SEQ ID NO. 10 or 11 and variants

FRL3	57 to 88	57	G
		58	V
		59	P
		60	D
		61	R
		62	F
		63	S
		64	G
		65	S
		66	G
		67	S
		68	G
		69	T
		70	D
		71	F
		72	T
		73	L
		74	K
		75	I
		76	S
		77	R
		78	V
		79	E
		80	A
		81	E
		82	D
		83	L or V
		84	G
		85	V
		86	Y
		87	Y

		88	C
CDRL3	89 to 97		SEQ ID NO. 12 or 13 and variants
FRL4	98 to 108	98	F
		99	G
		100	G or Q
		101	G
		102	T
		103	K
		104	L
		105	E
		106	I or L
		106a	K
		107	R
		108	A

The amino acid sequences SEQ ID Nos. 46 to 79 correspond to amino acid sequences with preferred framework sequences for the variable heavy chain. The amino acid sequences SEQ ID Nos. 80 to 94 correspond to amino acid sequences with preferred framework sequences for the variable light chain.

The techniques and methods to be used in the production of these sequences are well-known to those skilled in the art, and a person skilled in the art will be able to select suitable framework sequences and/or mutations.

In the meaning of the invention, core 1-specific recognition molecules can be present in different formats. The basic structure of the recognition molecule is one (or more) polypeptide chain(s) comprising the above-described inventive triplet sequence 1 or triplet sequences 1 and 2 and framework sequences. For example, the amino acid sequence

of the variable heavy chain is linked with the framework sequences and triplet sequences 1 and the amino acid sequence of the variable light chain is linked with the framework sequences and the triplet sequences 2 in a non-covalent or covalent fashion and can be situated on one or more polypeptide chains. A plurality of polypeptide chains can be present in covalently linked - e.g. via disulfide bridges - or non-covalently linked form as recognition molecule.

In particular, the various inventive formats of recognition molecules include linking of said triplet sequences with amino acid sequences beyond the framework sequences described above. In a preferred variant the recognition molecules according to the invention comprise further accessory sequences apart from the triplet sequences and framework sequences. More specifically, accessory sequences are amino acid sequences which primarily are not involved in the spatial configuration of the triplet sequences, such as in the form of framework sequences, but may have an advantageous influence thereon as a result of secondary or tertiary interactions. For example, accessory sequences in the form of constant domains of an antibody will stabilize the antibody, causing dimerization, thereby effecting improved binding of the antibody, or, for instance, fusion of an scFv with a domain of a bacteriophage coat protein causes an activity increase of scFv binding as disclosed in Jensen KB et al., 2002, for example.

In a preferred embodiment the recognition molecules comprise amino acid sequences with framework sequences on an antibody basis and further accessory sequences in addition to the triplet sequences. In particular, the accessory sequences assume at least one of the following functions:

a) linking a triplet sequence with its correspondingly suited framework sequences with at least one other triplet sequence with its correspondingly suited framework sequences in order to create or improve binding capability;

b) stabilization of domains, e.g. by means of a linker between two protein domains or amino acid sequences, which undergo interaction with others in the same or in a second chain;

c) effector functions for immunological purposes, e.g. by fusion with the Fc portion of antibodies, chemokines, cytokines, growth factors or parts thereof, or antibodies having a different specificity, or fragments thereof, for the recruitment of cells of the immune system, e.g. macrophages or parts of the complement system;

d) fusion with tags, e.g. multimerization sequences – for example, μ-tail sequence from IgM or association domain from p53 or MBL – for multimerization of the core 1-binding portions for multivalent binding or for purification of recognition molecules, e.g. His-tag, or for detection, e.g. myc-tag, or for labelling or chelating of recognition molecules e.g. by high-lysine sequences.

Suitable structures are well-known to those skilled in the art or can be derived from the prior art by logical deduction.

Further preferred embodiments are recognition molecules according to the invention comprising the following formats: single-chain antibody fragment (scFv), Fv fragment, Fab fragment, $F(ab)_2$ fragment, multibody (dia-, tria-, tetra-body), immunoglobulin of the IgG, IgM, IgA, IgE, IgD iso-

types or subclasses thereof, e.g. IgG1, or immunoglobulin-
derived recognition molecules comprising at least one con-
stant domain.

In a preferred embodiment the recognition molecules of the
invention are composed of a heavy and a light polypeptide
chain, each of the amino acid sequences of the heavy and
light chains comprising one of the above-described triplet
structures representing the CDR regions of the antibody,
the corresponding antibody framework sequences representing
the framework sequences of the antibody, and accessory se-
quences comprising at least one of the constant domains of
the antibody isotype. The two chains can form covalent
bonds with each other. The constant regions and variable
regions may include sequences of antibodies from one or
more species. Portions of constant domains or complete con-
stant domains can be deleted or mutated in order to e.g.
modify the effector function of accessory sequences, e.g.
to prevent or improve binding to Fc receptors. In a pre-
ferred embodiment the recognition molecule is a murine,
chimerized, humanized or human antibody or antibody frag-
ment. For example, chimerization is effected by linking the
variable antibody domains with constant antibody domains or
fragments of a constant domain of antibodies from different
species. Preferred are sequences of constant domains of hu-
man antibodies.

The antibody framework sequences can be selected in such a
way that the sequences are largely homologous to human an-
tibody sequences. Selection as to the species origin of the
framework sequences will also depend on the use. Thus, for
therapeutic use in particular fields, highest possible lev-
els of human framework sequences are preferred, particu-
larly in those cases where human anti-mouse antibody re-
sponse (HAMA) is to be avoided. In other therapeutic
fields, a xeno-portion is advantageous because it effects

additional stimulation of the immune system. A combination of both is particularly suitable in some cases, especially in those cases where a xeno-portion is advantageous in initial immunization and a species-compatible, i.e. a human portion, is advantageous in later uses.

Homology to human consensus sequences is preferred, with HuHI being preferred for the variable heavy chain, and HuKII being preferred for the variable light chain. Particularly preferred is homology to human germ line sequences which are known to those skilled in the art and can be obtained from the V BASE data base (www.mrc-cpe.cam.ac.uk), for example.

The techniques and methods to be used in the production of these sequences are well-known to those skilled in the art, and a person skilled in the art will also be able to select suitable human sequences and/or perform optionally required mutations of said sequences.

In another embodiment the triplet sequences generally corresponding to the binding loops (CDR regions) and preferably having high homologies to the corresponding sequence regions in the human germ line sequence are additionally adapted thereto step by step, using simple mutations, without impairing the specific binding to core 1. Recognition molecules having these sequences will be referred to as partially human antibodies or antibody fragments herein. For example, preferred humanized sequences are represented by the sequences SEQ ID Nos. 56 to 79 and SEQ ID Nos. 85 to 94, respectively.

In another preferred embodiment, specific amino acids of antibody framework sequences of a species are replaced by others in order to generate less immunogenic regions in general. This involves technologies *per se* known to those

skilled in the art, e.g. technologies of humanization, e.g.
CDR grafting, resurfacing, chain shuffling with mutations
and deimmunization by mutation or deletion of human MHC
epitopes.

In a preferred embodiment, this involves an IgM-derived
recognition molecule having the corresponding constant do-
mains of an IgM, preferably human sequences. In the meaning
of the invention, immunoglobulins are composed of a heavy
chain and a light chain of an antibody, and 2 light chains
and 2 heavy chains preferably represent a unit. Immu-
noglobulins of the IgM type usually consist of 5 such units
additionally linked via the J chain to form disulfide
bridges.

In a particularly preferred embodiment the J chain is ab-
sent, with multimerization of the subunits likewise taking
place, in which case hexa- and pentameric structures can be
present.

In a preferred embodiment of such recognition molecules,
single-chain antibody fragments are involved, comprising a
triplet structure 1 with the corresponding antibody frame-
work sequences described above, which represent the CDR re-
gions of the antibody and framework sequences of the vari-
able domain of the heavy chain of antibodies, and a triplet
structure 2 with the corresponding antibody framework se-
quences described above, which represent the CDR regions of
the antibody and framework sequences of the variable domain
of the light chain of antibodies, which are covalently
linked in the form of a fusion protein. Here, the sequences
are linked directly or via a linker. Preferred in this case
are scFv formats with no linker or with a linker 1 to 9
amino acids in length. The scFv antibodies form multimeric
structures (for example, dia-, tria-, tetrabodies) which,
in the meaning of the invention, are also referred to as

multibodies and exhibit higher avidity to the core 1 anti-
gen as a result of multivalence. Core 1-specific recogni-
tion molecules in an scFv format were constructed with
varying linker lengths (SEQ ID Nos. 95 to 106) and their
binding characteristics investigated in an ELISA. Step-by-
step linker length reduction resulted in an increase of
binding to asialoglycophorin, which is a core 1-bearing
glycoprotein, as illustrated in Figure 3. The variants hav-
ing SEQ ID Nos. 104 and 105 exhibited the best binding
properties. These multivalent constructs in a dia-/triabody
format are particularly preferred embodiments of the inven-
tion, being advantageous in tumor therapy as a result of
improved pharmacokinetic properties.

In another preferred embodiment the recognition molecules
are fused, chemically coupled, covalently or non-covalently
associated with (i) immunoglobulin domains of various spe-
cies, (ii) enzyme molecules, (iii) interaction domains,
(iv) signal sequences, (v) fluorescent dyes, (vi) toxins,
(vii) catalytic antibodies, (viii) one or more antibodies
or antibody fragments with different specificity, (ix) cy-
tolytic components, (x) immunomodulators, (xi) immunoeffec-
tors, (xii) MHC class I or class II antigens, (xiii) che-
lating agents for radioactive labelling, (xiv) radioiso-
topes, (xv) liposomes, (xvi) transmembrane domains, (xvii)
viruses and/or cells. In particular, the recognition mole-
cules can also be fused with a tag allowing detection of
the recognition molecule and purification thereof, such as
myc-tag or His-tag. Technologies for the production of
these constructs are well-known to those skilled in the
art, and a person skilled in the art will be able to select
suitable sequences and components and link them with the
recognition molecules of the invention in a suitable man-
ner.

In another preferred embodiment the above-described recognition molecules based on antibodies or antibody fragments are fused with peptides or proteins not derived from immunoglobulins. For example, the multimerization domain of a non-immunoglobulin molecule is fused with an scFv, especially the C-terminal end of the α-chain of the C4 binding protein, as described in Tonye Libyh M. et al., 1997, thereby constructing a multivalent recognition molecule.

In another embodiment, an scFv is fused with a transmembrane domain of a non-immunoglobulin molecule, e.g. with the transmembrane domain of c-erb B2, h-PDGFR, human transferrin receptor, or human asialoglycoprotein receptor (Liao et al., 2000), thereby enabling expression of binding molecules on the surface of cells.

Another preferred embodiment of the invention comprises recognition molecules according to the invention, additionally comprising amino acid sequences specifically binding to macrophages or other immunoeffector cells. For example, the recognition molecules of the invention further comprise an antibody binding site against CD64, and, in the form of a bispecific antibody or antibody fragment (diabodies), binding of macrophages to core 1-positive tumor cells takes place, resulting in combatting and/or destruction thereof.

A preferred embodiment of the invention relates to radiolabelled core 1-specific recognition molecules. One preferred form involves recognition molecules based on antibodies or antibody fragments. Another preferred embodiment involves radiolabelled recognition molecules of the invention in single-chain format (including the form of dia-, tria-, tetrabodies). Other preferred forms are radiolabelled single-chain antibody fragments and complete immunoglobulins, e.g. inventive chimeric or humanized IgG or IgM antibodies or humanized antibody fragments. It goes

without saying that the invention is not restricted to these antibodies, said radioactive labels and formats of antibodies.

Antibody fragments such as the preferred multivalent scFv fragments, especially with no or very short linker, offer an advantage in the targeting of solid tumors compared to intact monoclonal antibodies. With intact antibodies exhibiting specific accumulation within the tumor area in biodistribution studies, an inhomogeneous antibody distribution with primary accumulation in the peripheral regions is noted when precisely investigating the tumor. Due to tumor necroses, inhomogeneous antigen distribution and increased interstitial tissue pressure, it is not possible to reach central portions of the tumor with such antibody constructs. In contrast, smaller antibody fragments show rapid tumor labelling, penetrate deeper into the tumor, and also, are removed relatively rapidly from the bloodstream. However, the dissociation constant of monovalent antibody fragments such as Fabs or scFv frequently is excessively small, resulting in a short residence time on the tumor cells. For this reason, multivalent antibody constructs such as multibodies (diabodies, tria-/tetrabodies), $F(ab')_2$ and other minibodies (multivalent antibody constructs consisting of binding domain and multimerization sequence, e.g. scFv and CH3 domain of an IgG) offer many advantages in tumor therapy. Multivalent constructs in a dia-/triabody format are preferred embodiments of the invention, they are advantageous in tumor therapy as a result of improved pharmacokinetic properties and have been further developed for use in tumor therapy. They can be used as vehicles for specific accumulation of e.g. cytotoxic substances such as chemotherapeutic agents or radionuclides in a tumor. By suitably selecting the radionuclides, it is possible to destroy tumor cells over a distance of several cell diameters, so that even antigen-negative tumor cells in a tumor

area can be covered and poor penetration of antibodies into solid tumors can be compensated at least in part.

A particularly preferred embodiment of the invention involves radiolabelled multibodies - specifically as set forth in detail in Example 9 - which combine particularly advantageous pharmacokinetic properties and, in combination, have improved tumor retention, tumor penetration, serum half-life and serum to tumor distribution ratio compared to complete immunoglobulins and scFv. Further advantages are high avidity and bacterial expression, allowing low-cost production of such recognition molecules. Advantageously, this specific format of recognition molecules according to the invention is therefore suitable for use preferably in the treatment of small primary tumors, metastases and minimal residual diseases.

A preferred embodiment of the invention involves non-radiolabelled recognition molecules. One preferred form involves recognition molecules based on antibodies or antibody fragments.

A particularly preferred embodiment involves chimeric and humanized immunoglobulins based on IgM molecules for the inhibition of liver metastasization and control of residual tumor cells.

Other preferred embodiments are toxin- or cytostatic agent-coupled chimeric or humanized IgG- and IgM-based recognition molecules of the invention and, in particular, multibodies (dia-, tria-, tetrabodies) having particularly advantageous pharmacokinetic properties as set forth above.

Another preferred embodiment involves liposomes which are loaded with e.g. toxins or cytostatic agents and bear recognition molecules of the invention on the surface thereof.

A person skilled in the art will be able to select suitable radioisotopes, toxins and cytostatic agents. Suitable techniques, methods, dosages and formulations are well-known to those skilled in the art.

Another preferred embodiment of the invention involves effector cells of the immune system having recognition molecules of the invention bound on the surface thereof, which direct/address the effector cells to core 1-bearing tumor cells, thereby mediating control and/or destruction thereof. Preferred effector cells are macrophages, dendritic cells and NK cells obtained from the patient and coupled *ex vivo* with the recognition molecules. Also preferred are cell lines of these types of cells. Linking is effected e.g. by means of bispecific recognition molecules which, in addition to core 1-specific components, comprise amino acids which mediate binding to the effector cells. For example, these are bispecific antibodies, complement components or constant domains of antibodies.

Another preferred embodiment involves macrophages from a patient which, following collection, are coupled with a bispecific antibody, e.g. in the form of a complete antibody, preferably chemically coupled Fab fragments or, more preferably, diabodies which, on the one hand, recognize CD64 and, on the other hand, are core 1-specific according to the invention. These macrophages, which bear the bispecific recognition molecules via CD64 specificity, are re-administered to the patient in a suitable formulation in order to combat the core 1-positive tumor. The techniques used to this end, as well as suitable methods, dosages and formulations are well-known to those skilled in the art. Another preferred embodiment involves macrophages from a patient which, following collection, are coupled with a core 1-specific antibody or antibody fragment of the inven-

tion comprising the constant portion of an antibody which binds to macrophages via the *per se* known Fc receptors. The recognition molecules can bind to the macrophages either as complete antibodies, preferably chimeric or humanized IgG or IgM, or as antibody fragment, e.g. scFv, Fab or multibodies in the form of a fusion protein or chemically coupled with a portion of the constant domain of antibodies, which portion is well-known to those skilled in the art. The macrophages bearing the recognition molecules are re-administered to the patient in a suitable formulation in order to combat the core 1-positive tumor. The techniques used to this end, as well as suitable methods, dosages and formulations are well-known to those skilled in the art.

Another preferred embodiment involves cell lines or cells from the body, such as the above-described effector cells which are transfected with molecules comprising the core 1-specific recognition molecules of the invention and additional elements causing expression and anchoring in the membrane, e.g. transmembrane domain, and mediating activation of the effector cells upon contact with a core 1-bearing tumor cell. The appropriate elements are well-known to those skilled in the art. For example, a dendritic cell line is transfected with a vector comprising a recognition molecule which comprises an inventive scFv or multibody and a transmembrane domain and an activating domain. In another example, macrophages are virally transfected to this end. The effector cells bearing the recognition molecules are re-administered to the patient in a suitable formulation in order to combat the core 1-positive tumor. The techniques used to this end, as well as suitable methods, dosages and formulations are well-known to those skilled in the art.

The invention also relates to nucleic acid molecules comprising one or more genetic sequences which encode at least one of the above-described recognition molecules and/or

constructs according to the invention. Owing to the degenerate genetic code, said nucleic acid molecules may have highly varying sequences. The selection of the codon also depends on the cell used to produce the recognition molecules, because different codons frequently are preferred in different cells from different organisms, and there may be a strong influence on the expression rate; for example, the arginine codons AGA and AGG preferably utilized in eukaryotic genes are rarely seen in bacteria where the codons CGC and CGU are clearly more frequent. In preferred embodiments the nucleic acid molecule of the invention is a genomic DNA, a cDNA and/or an RNA. The criteria of selecting suitable codons and the production of a suitable nucleic acid molecule are well-known to those skilled in the art.

Furthermore, the invention relates to vectors for the expression of recognition molecules, specifically in cells. In the meaning of the invention, a vector is understood to be a nucleic acid molecule according to the invention, which serves to express the recognition molecule and comprises a nucleic acid sequence which includes one or more genetic sequences encoding at least one of the above-described recognition molecules and which, in particular, includes at least one promoter effecting expression of the recognition molecule. Of course, vectors may comprise additional elements well-known to those skilled in the art, which are used e.g. in the propagation of vectors for the production in suitable cells and in cloning. The nucleic acid sequences can be present on one or more vectors; in a preferred embodiment, for example, the heavy chain of an immunoglobulin of the invention is encoded by one and the light chain by another vector. In another preferred embodiment of the invention the variable domain of the light chain and the variable domain of the heavy chain are encoded as fusion protein on the same vector under one promoter. Furthermore, in the meaning of the invention, nu-

cleic acid sequences encoding portions of a recognition molecule can be expressed by different promoters well-known to those skilled in the art. In another embodiment, said different nucleic acid sequences can be present on one common vector. Each sequence can be expressed by its own - same or different - promoter, or the sequences can be present in a bicistronic vector under a promoter. In a preferred fashion, different expression rates of the components of recognition molecules are achieved by said different promoters, improving formation of the overall recognition molecule as compared to equal expression rate of different components. It is also preferred to use promoters which can be induced so as to improve expression of the recognition molecule. In a particularly preferred fashion the vectors also comprise the regulatory elements well-known to those skilled in the art, e.g. enhancers increasing expression of the recognition molecule or components thereof, e.g. the CMV enhancer or immunoglobulin enhancer sequences. The nucleic acid molecules and vectors preferably comprise additional nucleic acid sequences which are used as signal sequences for the secretion of recognition molecules or components thereof and are *per se* known to those skilled in the art, e.g. PelB, OmpA or MalE for prokaryotic cell systems, or the signal peptide of the T cell receptor, of immunoglobulin chains, of t-PA or EPO for eukaryotic cell systems [Boel et al., 2000; Herrera et al., 2000]. In an advantageous fashion, this facilitates the purification and/or improves the yield of recognition molecules. The methods for the production of the above-described nucleic acids and vectors, suitable promoters, enhancers and vector constructs, as well as the criteria for the selection thereof are well-known to those skilled in the art and will be explained in detail in the examples.

In a specific embodiment of the invention the vector according to the invention also comprises nucleic acid se-

quences encoding viral proteins. The virus itself will be referred to as one particular form of a vector, the genetic material of which comprises a nucleic acid sequence encoding a recognition molecule according to the invention. In a preferred form the recognition molecule is a fusion protein with a virus coat protein or components thereof, making it possible that not only the genetic material comprises the nucleic acid sequence of the recognition molecule, but also that the recognition molecule itself is present on the surface of the virus in a binding-active state, e.g. an scFv recognition molecule of the invention as a fusion protein with a coat protein of adenoviruses, poxviruses or vaccinia viruses suitable for gene-therapeutic uses. This mediates addressing the virus to a core 1-expressing tumor cell, so that expression of the recognition molecule in the tumor cell takes place. This can be utilized in the expression of the recognition molecule *in vivo* in the organism or *in vitro* in a cell culture. In a preferred fashion, well-known systems are employed which use a helper virus for replication so as to ensure the safety of a gene-therapeutic method comprising said vector. Methods for the production of the above-described viral vectors, for the infection and expression of recognition molecules are well-known to those skilled in the art.

In another specific embodiment the vector of the invention comprises a fusion protein of a recognition molecule according to the invention and a protein or peptide specifically binding to a virus. Advantageously, the recognition molecules obtained can be used to address the virus to a core 1-expressing cell. Thus, for example, transfer of the genetic material can be mediated via infections, thereby allowing expression of specific molecules - encoded by the genetic material of the virus - in cells *in vivo* in the organism in the form of a gene therapy or *in vitro* in a cell culture.

Furthermore, the invention relates to a method of obtaining said recognition molecules, comprising the incorporation of one or more vectors of the invention, which include one or more nucleic acid molecules of the invention, in a suitable host cell, culturing said host cell under suitable conditions, and providing one or more recognition molecules from the cells or from the culture medium. In the meaning of the invention, the term "incorporation of vectors" represents technologies *per se* known to those skilled in the art, by means of which said vector is introduced in a host cell, e.g. electroporation, transfection using cationic lipids or infection, remaining therein in a transient or stable fashion. In the meaning of the invention, the term "providing one or more recognition molecules" represents technologies *per se* known to those skilled in the art, by means of which the recognition molecules expressed during the culturing process are obtained from the culture supernatant and/or from the cells, e.g. various protein-chemical purification steps, e.g. fractionating, concentrating, precipitating and/or chromatography. The techniques and procedures to be used in this method are well-known to those skilled in the art, and a person skilled in the art will also be able to select suitable host cells and culturing conditions, as well as methods for the provision from cells and/or culture supernatants. For example, as set forth above, a person skilled in the art will select nucleic acid sequences with suitable codons and promoter sequences adapted to the host cell so as to obtain highest possible expression of active recognition molecules. In a preferred embodiment a person skilled in the art will use e.g. affinity-chromatographic steps, e.g. chromatography on protein A or protein G or protein L, or e.g. metal ion affinity chromatography via an additionally introduced His-tag. This will be illustrated in more detail in the examples.

Apart from the steps explicitly mentioned above, the term "obtaining" also comprises additional steps such as pre-treatment of the starting material or further treatments of the final product. Pretreatment procedures are *per se* known to those skilled in the art. In addition to the provision procedures described above, procedures of further treatment also comprise e.g. final composing and/or formulating the recognition molecule obtained by means of the production procedure into suitable forms of use and/or administration. The type of said forms of use and/or administration, e.g. solution, lyophilizate or tablet, will depend on the intended application. It is well-known to those skilled in the art which administration form is suitable for which purpose. Depending on the administration form, the recognition molecule produced using the method according to the invention can be present together with auxiliary agents, carriers or other active substances. Auxiliary agents are preferably adjuvants, other active substances, preferably immunostimulatory molecules such as interleukins. The recognition molecule produced using the method of the invention can also be chemically modified in further treatment steps. Preferably, the recognition molecule is suitably linked with one or more additional molecules, i.e. by chemical or physical interaction. As additional molecules in the meaning of the invention, other proteins or peptides are preferably used, which are covalently or non-covalently linked with the recognition molecule produced by means of the method according to the invention, e.g. in order to produce bispecific recognition molecules by linking a recognition molecule of the invention which specifically recognizes the core 1 antigen with a second molecule which e.g. specifically binds an immunoeffector cell (for example, macrophage, NK cells, dendritic cells), or e.g. a linkage with interleukins (for example, IL-2, IL-7, IL-12, IL-15), chemokines or growth factors, and by virtue of the effect of these molecules via binding of the recognition

molecule of the invention, immunoeffectors are directed to the core 1-positive tumor cells, combatting and/or destroying same, for example. As described above, said additional molecules or components thereof can also be part of the recognition molecule itself, in which case they would not be linked by means of the herein-described chemical or physical methods following expression of the recognition molecule. In the meaning of the invention, "immunoeffectors" are understood to be those components of the invention capable of directly or indirectly effecting control and/or destruction of core 1-positive tumor cells, e.g. immunoeffector cells such as macrophages, NK cells, dendritic cells, or effector molecules such as proteins or peptides of the complement system. Suitable as additional molecules within the scope of the method according to the invention are, in particular, substances developing a therapeutic or diagnostic effect, e.g. radioisotopes or toxins. These substances are linked with the recognition molecules using *per se* known procedures; for example, radioisotopes are either directly incorporated (for example, iodine) or bound via a covalently coupled chelating agent (for example, yttrium, indium, bismuth). The steps of the procedure of further treatment are well-known to those skilled in the art.

The cells used according to the invention to express the recognition molecules can be prokaryotic or eukaryotic cells, e.g. bacterial, yeast (preferably *S. cerevisiae* or *P. pastoris*), insect (*D. melanogaster*), plant, mammal cells (preferably hamster, mouse or human cell lines) or organisms such as transgenic animals and plants. Preferably, *E. coli* is used for expression of the recognition molecules of the invention in a prokaryotic system, and the mammal cell lines NS0, SP2/0, CHO-K1, CHOdhfr-, COS-1, COS-7, HEK293, K562, Namalwa or Percy 6 for expression in a eukaryotic system.

Furthermore, the present invention relates to host cells produced using the method described above, by means of which host cells recognition molecules of the invention can be produced. Of course, the host cells can be part of a clone or represent the clone themselves. The invention also relates to organisms comprising the host cells of the invention. Techniques to be used and methods of producing such organisms are well-known to those skilled in the art.

The invention also relates to compositions for therapeutic, prophylactic or diagnostic purposes, comprising at least one recognition molecule of the invention in a suitable, especially pharmaceutically suitable form or composition. More specifically, the pharmaceutical composition comprises additional materials and substances, e.g. medical and/or pharmaceutical-technical adjuvants. In the meaning of the invention, pharmaceutical compositions used for therapeutic and prophylactic purposes, as well as pharmaceutical compositions used as *in vivo* diagnostic agent will be regarded as drugs. In another preferred embodiment, compositions for *ex vivo* diagnostics are concerned, which may contain additional materials and substances. This embodiment will be illustrated in more detail in the description of diagnostic agents.

According to the invention, "drugs or pharmaceutical compositions" - used in a synonymous fashion herein - are substances and formulations of substances intended to cure, alleviate or avoid diseases, illness, physical defects or pathological affection by application on or in the human body. According to the invention, medical adjuvants are substances used as active ingredients in the production of drugs. Pharmaceutical-technical adjuvants serve to suitably formulate the drug or pharmaceutical composition and, if required during the production process only, can even be removed thereafter, or they can be part of the pharmaceuti-

cal composition as pharmaceutically tolerable carriers. Examples of pharmaceutically tolerable carriers will be given below. Drug formulation or formulation of the pharmaceutical composition is optionally effected in combination with a pharmaceutically tolerable carrier and/or diluent. Examples of suitable pharmaceutically tolerable carriers are well-known to those skilled in the art and include phosphate-buffered saline, water, emulsions such as oil/water emulsions, various types of detergents, sterile solutions, and so forth. Drugs or pharmaceutical compositions comprising such carriers can be formulated by means of well-known conventional methods. These drugs or pharmaceutical compositions can be administered to an individual at a suitable dose, e.g. in a range of from 1 µg to 10 g of recognition molecules per day and patient. Doses of from 1 mg to 1 g are preferred. Administration can be effected on various routes, e.g. intravenous, intraperitoneal, intrarectal, intragastrointestinal, intranodal, intramuscular, local, e.g. intratumoral, but also subcutaneous, intradermal or on the skin or via mucosa. Administration of nucleic acids can also be effected in the form of a gene therapy, e.g. by means of viral vectors described above. The kind of dosage and route of administration can be determined by the attending physician according to clinical factors. As is familiar to those skilled in the art, the kind of dosage will depend on various factors, such as size, body surface, age, sex, or general health condition of the patient, but also on the particular agent being administered, the time period and type of administration, and on other medications possibly administered in parallel.

More specifically, the pharmaceutical compositions or drugs comprise a pharmacological substance which includes one or more recognition molecules of the invention or/and nucleic acid molecules encoding same, in a suitable solution or administration form. Administration thereof can be effected

either alone or together with appropriate adjuvants described in connection with drugs or pharmaceutical compositions, or in combination with one or more adjuvants, e.g. QS-21, GPI-0100 or other saponines, water-oil emulsions such as Montanide adjuvants, polylysine, polyarginine compounds, DNA compounds such as CpG, Detox, bacterial vaccines such as typhoid vaccines or BCG vaccines and/or other suitable material enhancing the effect, preferably immunostimulatory molecules such as interleukins, e.g. IL-2, IL-12, IL-4 and/or growth factors such as GM-CSF. They are mixed with the recognition molecules of the invention according to well-known methods and administered in suitable formulations and dosages. Formulations, dosages and suitable components are well-known to those skilled in the art.

Obviously, the pharmaceutical composition or drug can also be a combination of two or more of the inventive pharmaceutical compositions or drugs, as well as a combination with other drugs, tumor vaccines or tumor treatments, such as antibody therapies, chemotherapies or radiotherapies, suitably administered or applied at the same time or separately in time. The production of the drugs or pharmaceutical compositions proceeds according to *per se* known methods.

In particular, the drugs or pharmaceutical compositions can be used in the treatment of core 1-positive tumor diseases such as mammary carcinomas, cervical carcinomas, ovarian carcinomas, colon carcinomas, gastrointestinal carcinomas, pancreas carcinomas, lung carcinomas, prostate carcinomas. Such tumor diseases may also include core 1- and/or core 2-positive tumor diseases. For example, the treatment is directed against primary tumors, minimal residual tumor diseases, relapses and/or metastases. The treatment of the tumors can also be effected as an adjuvant treatment. The drugs can also be used in the prophylaxis of core 1-

positive tumor diseases. For example, prophylactic use is directed to the prophylaxis of tumors and metastases. The tumor agents are administered in a suitable form according to well-known methods. A preferred variant is injection or administration of the drugs intravenously, locally in body cavities, e.g. intraperitoneal, intrarectal, intragastrointestinal routes, locally, e.g. directly in a tumor, in organs or lymphatic vessels (intranodal), but also subcutaneously, intradermally or on the skin, and intramuscularly. In a preferred fashion, types of administration can also be combined, in which case administration can be effected on different days of treatment or on one day of treatment. According to the invention, it is also possible to combine two or more of the inventive drugs or pharmaceutical compositions or one or more drugs of the invention with one or more drugs or tumor treatments, such as antibody therapies, chemotherapies or radiotherapies, suitably administered or applied at the same time or separately in time.

The present invention also relates to a method for the production of a drug or a pharmaceutical composition, comprising the steps of producing recognition molecules and further comprising the step of formulating the recognition molecules of the invention into a pharmaceutically tolerable form. The recognition molecules preferred to this end are described above as further embodiments of the treatment of tumor diseases and prophylaxis, as well as under *in vivo* diagnostic agents below.

Hence, the recognition molecules of the invention and the substances and compositions produced using the method according to the invention can be used in a preferred fashion in prophylaxis, diagnosis, follow-up and/or treatment of tumor diseases. Furthermore, it is preferred to use the recognition molecules, vectors and/or the drug or pharma-

ceutical composition in the prophylaxis and/or treatment of cancer diseases, including tumors and metastases.

In a preferred embodiment the cancerous disease or tumor being treated or prevented is selected from the group of cancerous diseases or tumor diseases of the ear-nose-throat region, of the lungs, mediastinum, gastrointestinal tract, urogenital system, gynecological system, breast, endocrine system, skin, bone and soft-tissue sarcomas, mesotheliomas, melanomas, neoplasms of the central nervous system, cancerous diseases or tumor diseases during infancy, lymphomas, leukemias, paraneoplastic syndromes, metastases with unknown primary tumor (CUP syndrome), peritoneal carcinomatoses, immunosuppression-related malignancies and/or tumor metastases.

More specifically, the tumors may comprise the following types of cancer: adenocarcinoma of breast, prostate and colon; all forms of lung cancer starting in the bronchial tube; bone marrow cancer, melanoma, hepatoma, neuroblastoma; papilloma; apudoma, choristoma, branchioma; malignant carcinoid syndrome; carcinoid heart disease, carcinoma (for example, Walker carcinoma, basal cell carcinoma, squamobasal carcinoma, Brown-Pearce carcinoma, ductal carcinoma, Ehrlich tumor, *in situ* carcinoma, cancer-2 carcinoma, Merkel cell carcinoma, mucous cancer, non-parvicellular bronchial carcinoma, oat-cell carcinoma, papillary carcinoma, scirrhus carcinoma, bronchio-alveolar carcinoma, bronchial carcinoma, squamous cell carcinoma and transitional cell carcinoma); histiocytic functional disorder; leukemia (e.g. in connection with B cell leukemia, mixed-cell leukemia, null cell leukemia, T cell leukemia, chronic T cell leukemia, HTLV-II-associated leukemia, acute lymphocytic leukemia, chronic lymphocytic leukemia, mast cell leukemia, and myeloid leukemia); malignant histiocytosis, Hodgkin disease, non-Hodgkin lymphoma, solitary plasma cell

tumor; reticuloendotheliosis, chondroblastoma; chondroma, chondrosarcoma; fibroma; fibrosarcoma; giant cell tumors; histiocytoma; lipoma; liposarcoma; leukosarcoma; mesothelioma; myxoma; myxosarcoma; osteoma; osteosarcoma; Ewing sarcoma; synovioma; adenofibroma; adenolymphoma; carcinosarcoma, chordoma, craniopharyngioma, dysgerminoma, hamartoma; mesenchymoma; mesonephroma, myosarcoma, ameloblastoma, cementoma; odontoma; teratoma; thymoma, chorioblastoma; adenocarcinoma, adenoma; cholangioma; cholesteatoma; cylindroma; cystadenocarcinoma, cystadenoma; granulosa cell tumor; gynadroblastoma; hidradenoma; islet-cell tumor; Leydig cell tumor; papilloma; Sertoli cell tumor, theca cell tumor, leiomyoma; leiomyosarcoma; myoblastoma; myoma; myosarcoma; rhabdomyoma; rhabdomyosarcoma; ependymoma; ganglioneuroma, glioma; medulloblastoma, meningioma; neurilemmoma; neuroblastoma; neuroepithelioma, neurofibroma, neuroma, paraganglioma, non-chromaffin paraganglioma, angiokeratoma, angiolymphoid hyperplasia with eosinophilia; sclerotizing angioma; angiomatosis; glomangioma; hemangioendothelioma; hemangioma; hemangiopericytoma, hemangiosarcoma; lymphangioma, lymphangiomyoma, lymphangiosarcoma; pinealoma; cystosarcoma phylloides; hemangiosarcoma; lymphangiosarcoma; myxosarcoma, ovarian carcinoma; sarcoma (for example, Ewing sarcoma, experimentally, Kaposi sarcoma and mast cell sarcoma); neoplasms (for example, bone neoplasms, breast neoplasms, neoplasms of the digestive system, colorectal neoplasms, liver neoplasms, pancreas neoplasms, hypophysis neoplasms, testicle neoplasms, orbital neoplasms, neoplasms of the head and neck, of the central nervous system, neoplasms of the hearing organ, pelvis, respiratory tract and urogenital tract); neurofibromatosis and cervical squamous cell dysplasia.

In another preferred embodiment the cancerous disease or tumor being treated or prevented is selected from the group of cancerous diseases or tumor diseases comprising cells

including the core 1 in the definition according to the invention, selected from the group of: tumors of the ear-nose-throat region, comprising tumors of the inner nose, nasal sinus, nasopharynx, lips, oral cavity, oropharynx, larynx, hypopharynx, ear, salivary glands, and paragangliomas, tumors of the lungs, comprising non-parvicellular bronchial carcinomas, parvicellular bronchial carcinomas, tumors of the mediastinum, tumors of the gastrointestinal tract, comprising tumors of the esophagus, stomach, pancreas, liver, gallbladder and biliary tract, small intestine, colon and rectal carcinomas and anal carcinomas, urogenital tumors comprising tumors of the kidneys, ureter, bladder, prostate gland, urethra, penis and testicles, gynecological tumors comprising tumors of the cervix, vagina, vulva, uterine cancer, malignant trophoblast disease, ovarian carcinoma, tumors of the uterine tube (Tuba Faloppii), tumors of the abdominal cavity, mammary carcinomas, tumors of the endocrine organs, comprising tumors of the thyroid, parathyroid, adrenal cortex, endocrine pancreas tumors, carcinoid tumors and carcinoid syndrome, multiple endocrine neoplasias, bone and soft-tissue sarcomas, mesotheliomas, skin tumors, melanomas comprising cutaneous and intraocular melanomas, tumors of the central nervous system, tumors during infancy, comprising retinoblastoma, Wilms tumor, neurofibromatosis, neuroblastoma, Ewing sarcoma tumor family, rhabdomyosarcoma, lymphomas comprising non-Hodgkin lymphomas, cutaneous T cell lymphomas, primary lymphomas of the central nervous system, Hodgkin's disease, leukemias comprising acute leukemias, chronic myeloid and lymphatic leukemias, plasma cell neoplasms, myelodysplasia syndromes, paraneoplastic syndromes, metastases with unknown primary tumor (CUP syndrome), peritoneal carcinomatosis, immunosuppression-related malignancy comprising AIDS-related malignancies such as Kaposi sarcoma, AIDS-associated lymphomas, AIDS-associated lymphomas of the central nervous system, AIDS-associated Hodgkin disease, and AIDS-associated ano-

genital tumors, transplantation-related malignancy, metastasized tumors comprising brain metastases, lung metastases, liver metastases, bone metastases, pleural and pericardial metastases, and malignant ascites.

In another preferred embodiment the cancerous disease or tumor being treated or prevented is selected from the group comprising cancerous diseases or tumor diseases such as mammary carcinomas, gastrointestinal tumors, including colon carcinomas, stomach carcinomas, pancreas carcinomas, colon cancer, small intestine cancer, ovarian carcinomas, cervical carcinomas, lung cancer, prostate cancer, renal cell carcinomas and/or liver metastases.

The recognition molecules of the invention can be directly employed in the treatment or prophylaxis of tumor diseases or coupled with additional effector structures. According to the invention, "effector structures" are understood to be chemical or biochemical compounds, molecules or atoms which directly or indirectly cause destruction or damage, including e.g. growth reduction or growth inhibition, of tumor cells. For example, these include radioisotopes, toxins, cytostatic agents and other effector molecules such as cytokines and chemokines or other structures representing effectors themselves or being coupled to said effector molecules, e.g. liposomes loaded with toxins or cytostatic agents, which bear the recognition molecules according to the invention. In the latter example of liposomes, particularly those effector structures are concerned which, in addition to the recognition molecule for tumor specificity, bear molecules responsible for reception of effector structures or components thereof in cells, such as antibodies against receptors causing receptor-mediated endocytosis. In such cases, the recognition molecules preferably comprise a transmembrane domain allowing their insertion in the liposomal membrane, or, in another preferred embodiment the

recognition molecules are chemically coupled on the liposome surface. The techniques used to this end are well-known to those skilled in the art, including production of the liposomes. Linking of the recognition molecules with other effector structures also proceeds according to *per se* known methods. As already set forth above, linking can be effected e.g. directly by covalent or non-covalent loading, by chemical coupling, which may require an additional chemical or biological molecule, e.g. a chelating agent or linker, or in the form of fusion proteins or peptides via fusion. The recognition molecules are employed in the treatment of tumor diseases with core 1-bearing tumors and/or - for a subgroup of recognition molecules of the invention described above for their specificity for core 1 and core 2 - core 2 and/or core 1-bearing tumor cells or in prophylaxis which, for example, prevents formation of primary tumors or metastases. One preferred objective is treatment of minimal residual disease and of metastases. Another preferred use is inhibition of liver metastasization of core 1 and/or core 2-positive tumor cells. The recognition molecules according to the invention are administered in a suitable formulation, in one go or repeatedly, at suitable intervals and in suitable doses.

Infra and *supra*, in the meaning of the invention the core 1 antigen is understood to be also core 1 and/or core 2, and core 1-positive cells or tumor cells and/or tissues are understood to be also core 1 and/or core 2-positive cells or tumor cells and/or tissues.

In a preferred embodiment the above-described radioactive recognition molecules according to the invention are combined with an application of non-labelled core 1-specific recognition molecules according to the invention. This helps towards an improvement of the background and more specific binding to the tumor by saturating potential core

1-bearing molecules in the blood. To this end, IgM-derived recognition molecules are preferably used, e.g. the cIgM described in the examples or a humanized form thereof, because they primarily bind to core 1 antigen in blood, thereby reducing the background and serum radioactivity load and increasing the relative tumor targeting, while limiting penetration into tissues and tumors by virtue of the size of the molecules. The procedures and technologies used to this end are well-known to those skilled in the art, and a person skilled in the art will also be able to devise a suitable dose, formulations, route of application, and time of administering said non-labelled recognition molecules.

Also preferred is the use of viral vectors in gene-therapeutic applications wherein specifically the surface of the viruses bears recognition molecules according to the invention.

The invention also relates to methods using the recognition molecules according to the invention, which methods allow identification and/or recovery of core 1-bearing molecules from a large pool of different molecules, which can be used with advantage in applications in tumor treatment, tumor prophylaxis and tumor diagnosis. According to the invention, core 1-bearing molecules are understood to be molecules which bear core 1 and/or core 2 structures and are bound by the recognition molecules of the invention in a specific fashion. According to the invention, core 1-bearing molecules are glycoproteins, glycopeptides and/or glycolipids, as well as cells or other vehicles, such as viruses, bacteria, components of cells, such as exosomes or cell lysates, or liposomes, which contain one or more core 1 structures. The core 1-bearing molecules can be accumulated or isolated from cells or cell lines, culture super-

natants, tumor tissues, tumor cells, or body fluids such as blood, blood serum, lymph, urine, spinal fluid or sperm.

Mutatis mutandis, the definitions of terms introduced above also apply to terms in the methods described below.

Core 1-bearing molecules are identified and/or isolated and obtained in a method of the invention by binding to the above-described core 1-specific recognition molecules according to the invention. According to the method of the invention, the above-described core 1-bearing molecules can be obtained from body fluids or from supernatants of cell cultures by means of affinity chromatography. It is possible to combine further purification and/or concentration steps with one or more affinity-chromatographic steps according to *per se* known methods. Likewise, tumor-associated core 1-bearing molecules can be obtained from tumor cells, tumor tissues or tumor cell lines by upstream insertion of a suitable step according to *per se* known methods, so that cell-associated core 1-bearing molecules can be put to affinity purification, e.g. by solubilization with suitable detergents or by cleavage using proteolysis or by cell lysis.

In another method of the invention, core 1-bearing molecules or cells are obtained from tissues. To this end, the tissue is digested according to *per se* known methods in order to provide access to the core 1-bearing molecules or cells, e.g. by means of proteolytic or mechanical methods. Such methods are well-known to those skilled in the art.

As set forth above, core 1-positive cells or cell lines are also isolated or accumulated using said core 1-specific recognition molecules and separated from cells bearing no or low quantities of core 1 structures. The term "isolation or accumulation of cells" is understood to include all

measures of separating cells having formed a complex with the recognition molecules of the invention as a result of bearing said core 1 structures. Such methods are well-known to those skilled in the art. In a preferred fashion, FACS or MACS methods are employed to this end. For example, accumulation proceeds via binding of recognition molecules of the invention to the core 1 structure on the cell surface and subsequent selection of thus labelled cells by binding to carrier materials specifically interacting with the recognition molecule, e.g. anti-mouse IgM antibodies coupled to magnetic beads (MAC sorting). Furthermore, the core 1-specific recognition molecules can be coupled covalently to a carrier. Another example is recovery using an FAC sorter which sorts cells bearing fluorescence-labelled recognition molecules. Both of these methods are well-known to those skilled in the art. The core 1-positive cells accumulated in this way can be used in the production of vaccines, e.g. for loading dendritic cells or directly as tumor cell lysate in a vaccine composition. Previous accumulation of core 1-positive cells is to provide higher tumor specificity of vaccination. These methods are well-known to those skilled in the art.

The present invention also relates to methods of producing a diagnostic agent, comprising the steps of the inventive method for the production of core 1-specific recognition molecules according to the invention and, in addition, comprising the step of formulating the recognition molecules in a diagnostically usable form.

According to the invention, the term "diagnostic agent" defines substances and preparations of substances intended to recognize diseases, illness, physical defects or pathological affection by application on or in the human body. Preferably, parts of the human body are understood to be body

fluids such as blood, blood serum, lymph, urine, spinal
fluid or sperm, or tissue biopsies or samples.

Formulating the diagnostic agent preferably comprises modi-
fication of the produced recognition molecules with sub-
stances allowing detection of the core 1 antigen and also,
in specific embodiments depending on the fine specificity
of the recognition molecule according to the invention, of
core 2 antigen by definition. Suitable substances are well-
known in the art. Based on the selection of a substance, a
person skilled in the art will be able to take suitable
measures in order to formulate a diagnostic agent.

According to the invention, it is also possible for diag-
nostic purposes to couple substances to the recognition
molecules according to *per se* known methods, which facili-
tate detection of core 1 antigens and/or carrier molecules
and/or cells thereof, e.g. by biotinylation, fluorescence
labelling, radioactive labelling or enzyme linking of rec-
ognition molecules.

Another method of tumor diagnostics and prognosis uses rec-
ognition molecules of the invention which recognize core 1
antigens and/or carrier molecules thereof in serum of hu-
mans. Determination is preferably qualitative, quantitative
and/or in time-dependent relative quantities according to
per se known methods. According to the invention, the same
methods are also used in the follow-up of tumor diseases
and to control the course of treatment, including monitor-
ing of immune responses, and for control and dosage of tu-
mor treatments. The techniques used in such methods are *per
se* well-known, e.g. ELISA, Western blot, FACS (fluores-
cence-activated cell sorting), MACS (magnetic-activated
cell sorting), ADCC (antibody-dependent cell cytotoxicity),
CDC (complement-dependent cytotoxicity), immunocytochemis-
try and immunohistochemistry.

The preferred inventive methods of tumor diagnostics and prognosis use core 1-specific recognition molecules of the invention in *per se* well-known methods to detect the core 1 antigen in serum or in tissue preparations. In these methods, core 1 antigen on carrier molecules, core 1 present in immune complexes on carrier molecules and/or core 1 bound on cells is detected, and the presence of core 1 antigen and/or core 1-bearing molecules is determined qualitatively, quantitatively and/or in relative quantities according to *per se* known methods. According to the invention, the same methods are employed in the follow-up of tumor diseases and to control the course of treatments. The techniques used in such methods are *per se* well-known, e.g. ELISA, Western blot, FACS (fluorescence-activated cell sorting), MACS (magnetic-activated cell sorting), ADCC (antibody-dependent cell cytotoxicity), CDC (complement-dependent cytotoxicity), immunocytochemistry and immunohistochemistry.

One preferred embodiment is a tissue rapid test wherein the tissue samples are stained with fluorescence-labelled recognition molecules of the invention in a immunohistological method. In another preferred method the recognition molecule according to the invention, preferably an isotype IgM antibody, is combined with another antibody specifically recognizing the MUC1 antigen, preferably isotype IgG1. The advantage is that, e.g. in gastrointestinal carcinoma diagnostics (e.g. colorectal carcinomas and stomach carcinomas), recognition at an early stage and, at the same time, prognosis with respect to the course of disease and/or risk of liver metastasization is possible, higher levels of core 1 antigen indicating a more unfavorable prognosis as to the course and a probability of liver metastasization increased by several times. In another preferred embodiment the antibodies and recognition molecules are directly labelled with various fluorescent dyes, e.g. Cy3 and Cy5 or Cy3 and FITC.

In one embodiment, wherein signal intensification is advantageous, the antibodies and/or recognition molecules are enhanced by labelled secondary antibodies or biotin-streptavidin. Advantageously, different isotypes and/or sequences of species in the constant region of antibodies are used. The techniques and methods used to this end, e.g. of labelling and immunohistology, as well as the selection of suitable formats of recognition molecules are well-known to those skilled in the art. The diagnostic method described above is not restricted to gastrointestinal tumors, but can be used in any tumor disease involving the core 1 antigen.

In another preferred embodiment a serological test is performed, using a sandwich ELISA procedure. This consists of a scavenger antibody, which binds carrier molecules of the core 1 antigen from serum to a solid phase, and a detection antibody which, according to the invention, also includes other recognition molecules of the invention which recognize the core 1 antigen. In this way, it is possible to distinguish which carrier molecule is the one that bears core 1. In a preferred form it is possible to draw conclusions about the origin of the primary tumor. A variety of antibodies recognizing glycoproteins bearing O-glycosylations can be used as scavenger antibodies. A preferred embodiment uses antibodies against the MUC1 epithelial mucin as scavenger antibody, which frequently bears core 1 in tumor cases. In another embodiment, all antigens bearing the core 1 antigen are determined in blood. This is possible because the core 1 antigen is present in a plurality of copies per carrier molecule. According to the invention, a core 1-specific recognition molecule of the invention is used as scavenger antibody, and a labelled core 1-specific recognition molecule of the invention is used as detection antibody, in which case the recognition molecules do not have to be antibodies. In a preferred embodiment an IgM as recognition molecule is used at least as scavenger or de-

tection antibody. In another preferred embodiment the de-
tection antibody is labelled with biotin, and the system is
detected via streptavidin in combination with a suitable
detection method. For example, suitable detection methods
are POD labelling or fluorescence labelling of strepta-
vidin.

For a serological tumor test, another preferred embodiment
of the invention combines the determination of core 1, as
described above, with the determination of other serologi-
cal tumor markers, e.g. PSA, CEA or AFP. One embodiment
preferred in this case is determination of MUC1 and core 1
antigen. In a preferred embodiment, MUC1 is immobilized
from the serum on a solid phase, using an MUC1-specific an-
tibody, and detected with a second, anti-MUC1-specific an-
tibody as detection antibody, preferably one with improved
recognition of the DTR region in glycosylated form, and the
core 1 antigen is detected on MUC1 immobilized by means of
an anti-MUC1 scavenger antibody, using a recognition mole-
cule according to the invention. This diagnostic test com-
bines early recognition with a prognostic statement as to
the course of disease and/or the probability of liver me-
tastasization. The techniques used to this end, e.g. label-
ling and serology, including the detection methods, are
well-known to those skilled in the art. The diagnostic
methods described above are not restricted to gastrointes-
tinal tumors, but can be used in any tumor bearing the core
1 antigen. The serological tests described above are used
in diagnosis, monitoring the course of a tumor disease, and
in the prognosis of core 1 antigen-positive tumors.

In another method according to the invention, the core 1-
specific recognition molecules of the invention are used in
in vivo diagnostics. To this end, the recognition molecules
are labelled using suitable, *per se* known methods and thus
made available for *per se* known imaging methods in humans,

e.g. radioimmunodiagnostics, PET scanning methods or immunofluorescence endoscopy, e.g. by coupling and/or loading with appropriate molecules, e.g. radioactive isotopes such as indium, or fluorescent dyes such as Cy3, Cy2, Cy5 or FITC. In a preferred embodiment, multibodies according to the invention are covalently coupled with a suitable chelating agent (for example, DOTA or DTPA) and, loaded with indium-111, used in *in vivo* diagnostics. In a preferred embodiment, they are administered intravenously at a dose appropriate to the individual, and the location of the core 1 antigen and of a potential tumor is measured according to *per se* known methods. The methods and technologies used to this end, including imaging methods, are well-known to those skilled in the art, and a person skilled in the art will also be able to devise a suitable dose and formulations.

In another preferred embodiment, immunoglobulins, preferably IgM and IgG, are radiolabelled as described above and illustrated in more detail in the examples, e.g. with indium-111, and administered locally into the tumor or blood vessels supplying or evacuating the tumor. In one embodiment, this is used to determine the size of the tumor, and in another embodiment, to determine affected lymphatic nodes. The methods and technologies used to this end are well-known to those skilled in the art, and a person skilled in the art will also be able to devise a suitable dose and formulations.

In another embodiment the radioactively labelled recognition molecules are also administered via other routes of application. Preferred routes are intraperitoneal, intranodal or intrarectal and intragastrointestinal, respectively. Intraperitoneal is particularly advantageous in the determination of tumors accessible through the peritoneum and/or metastasizing therein, e.g. ovarian carcinomas and

certain gastrointestinal carcinomas. Intrarectal or intra-
gastrointestinal administration is advantageous in some
gastrointestinal tumors and in localization and size deter-
mination thereof. In some cases, intranodal can be used for
direct infiltration of single lymphatic nodes.

In a preferred embodiment the above-described radioactive
recognition molecules are combined with an application of
non-labelled core 1-specific recognition molecules of the
invention for *in vivo* diagnostic agents. This is to improve
the background. To this end, IgM-derived recognition mole-
cules are preferably used because they primarily bind to
core 1 antigen in blood, thereby significantly reducing the
background, while limiting penetration into tissues and tu-
mors by virtue of the size of the molecules. The methods
and technologies used to this end are well-known to those
skilled in the art, and a person skilled in the art will
also be able to devise a suitable dose, formulations, route
of application, and time of administering said non-labelled
recognition molecules.

In another preferred embodiment, recognition molecules of
the invention, preferably immunoglobulins, multibodies or
antibody fragments, more preferably IgM, IgG and multibod-
ies, are labelled with a fluorescent dye and administered
in vivo. Preferred routes of application are intrarectal,
intragastrointestinal, intraperitoneal, intravenous and
into supplying or evacuating blood vessels. A particularly
preferred embodiment is used to localize gastrointestinal
carcinomas by means of fluorescence endoscopy following ap-
plication of fluorescence-labelled recognition molecules.
In another preferred embodiment a recognition molecule of
the invention is combined with at least one antibody to an-
other tumor antigen, preferably anti-MUC1 antibody. In a
preferred fashion, different fluorescent dyes are used, al-
lowing differentiation of the recognition molecules and an-

tibodies, thereby combining a prognostic statement with early recognition and a greater number of cases. Preferred fluorescent dyes are those having lower background fluorescence, which are well-known to those skilled in the art. The methods and technologies used to this end, including imaging methods, e.g. fluorescence endoscopy, are well-known to those skilled in the art, and a person skilled in the art will also be able to devise a suitable dose, formulations, route of application, and time of administering said non-labelled recognition molecules.

The invention has several advantages: The core 1-specific recognition molecules of the invention recognize the types of carcinomas in a specific fashion, which is why they can be used with advantage in diagnosis and/or therapy of a large number of tumor patients with different indication. Moreover, the recognition molecules advantageously show virtually no binding on normal tissues. Compared to well-known tumor markers, this is a particular advantage and an outstanding property of the recognition molecules according to the invention. Another advantage is that the recognition molecules recognize the core 1 antigen independently of the carrier. One particular advantage of the recognition molecules of the invention is their high specificity for tumor tissue. In particular, this is due to the high specificity for definite carbohydrate antigens. Namely, non-specific recognition of other carbohydrate structures would increase the risk of non-specific recognition of non-tumor tissue. Furthermore, the recognition molecules of the invention exhibit high affinity. In particular, this presents a way of constructing lower-valent fragments such as IgG and multibodies. The option of having these different formats available is advantageous in the development of therapeutic agents. The core 1 and/or core 2 structures on the cell surface increase the probability of metastase formation, e.g. of liver metastases; by blocking the core 1 and/or

core 2 structures with recognition molecules, formation of metastases is reduced or inhibited.

Without intending to be limiting, the invention will be explained in more detail with reference to the examples.

Examples

1. Preparation of core 1-specific multibodies with short linkers

Multibodies having the sequences SEQ ID Nos. 96 to 106 were formed by shortening or deletion of the linker between the V_H and V_L of the single-chain antibody having the sequence SEQ ID NO. 95 (Fig. 1a). To this end, V_H and V_L were amplified with specific primers in such a way that 22 nucleotides at the 3' end of V_H and at the 5' end of V_L formed a complementary region (Fig. 1b, PCR I and PCR II) and subsequently, following purification, the two PCR fragments were linked in an SOE-PCR (Fig. 1b, PCR III). Finally, the PCR fragment was cloned into a prokaryotic expression vector via NcoI/NotI. This vector includes the lacZ promoter, a ribosome binding site (RBS), the M13 origin, the pelB signal sequence for secretion into the periplasm, an ampicillin resistance gene, and a cloning cassette to couple a hexahistidine tag for efficient purification and a c-myc-tag to the C-terminal end of the scFv (Fig. 2).

2. Bacterial expression and purification of the core 1-specific multibodies

The antibody fragments from Example 1 were expressed in *Escherichia coli* and purified. To this end, the corresponding plasmid was transformed in electrocompetent *E. coli* by means of electroporation and cultured in 2xTY medium (10 g

of yeast extract, 16 g of tryptone, 5 g of NaCl per liter) with 100 µg/ml ampicillin overnight. This culture was diluted 1:100 with 2xTY medium added with 100 µg/ml ampicillin and 0.5% glucose and incubated at 37°C until an $OD_{600 \, nm}$ of about 0.6 was reached. Thereafter, the culture was added with 1 mM IPTG for induction and incubated at 25°C for another 5 hours. The bacteria were harvested by centrifugation at 4000 × g for 20 min, the cell pellet was resuspended in TES buffer (30 mM Tris-HCl, pH 8.0, 20% saccharose, 1 mM EDTA) and incubated on ice for 20 min. Subsequently, 5 mM $MgSO_4$ was added, and the suspension was incubated on ice for another 20 min. The periplasm fraction was obtained by centrifugation at 4000 × g for 60 min and dialyzed against binding buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl, 10 mM imidazole) at 4°C overnight. The antibody fragments contained in the periplasm fraction were purified by metal ion affinity chromatography (HiTrap Chelating HP, Amersham Pharmacia Biotech) using the C-terminal His-tag. To this end, the dialyzed fraction was loaded on a column previously equilibrated with binding buffer, and the non-binding proteins were washed from the column with washing buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl, 30 mM imidazole). Subsequently, the antibody fragments were eluted with elution buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl, 300 mM imidazole). The above purification protocol was used for all core 1-specific antibody fragments having a hexahistidine tag, e.g. the humanized single-chain antibodies from Example 6.

3. **Analysis of core 1-specific multibodies in scFv format with varying linker length in an ELISA**

Multibodies having the amino acid sequences SEQ ID Nos. 95, 96, 97, 98, 99, 100, 101, 103, 104 and 105 were expressed in *E. coli* as described above and the periplasm fractions obtained. Asialoglycophorin (Sigma), which is a core 1-

bearing glycoprotein, was used as antigen in the ELISA. Using stock solutions (1 mg in 1 ml of bidist. H_2O) stored in portions at -20°C, a dilution of 5 µg/ml in PBS was produced. 50 µl/well of the above was pipetted in a microtiter plate (NUNCLON-TC Microwell 96 F), and the test plate was incubated at 4°C overnight. On the next day, the test plate was washed 3 times with PBS/0.2% Tween. Subsequently, non-specific binding sites were blocked with 2% BSA in PBS, and 50 µl of each fraction diluted with PBS/1% BSA in different dilution steps was applied and incubated at 37°C for 2 hours. After three wash steps with PBS/0.2% Tween, peroxidase-coupled anti-His-tag antibodies were employed as secondary antibodies to detect the specifically bound antibody constructs. To detect the bound secondary antibody, a color reaction with TMB (3,3',5,5'-tetramethylbenzidine) was performed. After 15 minutes the reaction was quenched by adding 2.5 N H_2SO_4. Measurement was performed using a microtiter plate photometer with 450 nm filter in dual mode versus 630 nm reference filter. The result is illustrated in Figure 3. Step-by-step linker length reduction results in increased binding to asialoglycophorin. The best binding properties are seen in the variants having SEQ ID Nos. 104 and 105. These multivalent constructs in dia/triabody format are preferred embodiments of the invention and offer advantages in tumor therapy owing to their improved pharmacokinetic properties.

4. Cloning of vectors to express chimeric core 1-specific IgG and IgM antibodies

The NcoI/XhoI DNA fragment from the scFv vector, which encodes V_H (Fig. 4), was cloned into the NcoI/SalI-cut BS Leader vector. The BS Leader vector includes a cloning cassette to introduce the T cell receptor signal peptide sequence at the 5' end and a splice donor sequence at the 3' end of the sequences of the variable domains (Fig. 4). The

V_L sequence of the corresponding antibody was amplified with specific primers to introduce the NcoI restriction site at the 5' end and the NheI restriction site at the 3' end in the PCR using the scFv sequence as template and, following NcoI/NheI digestion, cloned into the likewise digested BS Leader vector. Thereafter, each HindIII/BamHI fragment from the BS Leader vector was cloned into the corresponding eukaryotic expression vector. These vectors (pEFpuroCγ1V_H, pEFpuroCµV_H and pEFneoCκV_L) include EF-1α-promoter and HCMV enhancer, SV40 origin, BGH polyadenylation signal, puromycin resistance gene in the vector for the heavy chain and neomycin resistance gene or dehydrofolate reductase gene in the vector for the light chain, as well as the genomic sequences of the human constant γ1 region or µ region for the heavy chain or of the human constant κ region for the light chain (primers for amplification from genomic human DNA and vector map see Fig. 4).

5. Eukaryotic expression of core 1-specific chimeric IgG and IgM antibodies in CHO cells and purification thereof

To express the chimeric antibodies cIgG-Karo4 consisting of the sequences SEQ ID Nos. 111 and 113 and cIgM-Karo4 consisting of the sequences SEQ ID Nos. 112 and 113, CHOdhfr-cells (ATCC No. CRL-9096) were co-transfected with a mixture of vectors for the heavy and light chains (1:3) by means of electroporation (10^6 cells/ml, 500 V, 50 s) and cultured in selection medium (CHO-S-SFM II medium (Life Technologies), HT supplement (Biochrom), 400 µg/ml G418, 5 µg/ml puromycin) for 2 weeks. Following single-cell cloning in a 96-well plate, the supernatants were tested in an ELISA (asialoglycophorin as antigen, anti-human Fcγ1-POD-coupled or anti-human Fc5µ-POD-coupled (Dianova) as secondary antibody), and the clone with the highest antibody production rate was selected (about 0.5 µg/10^6 cells/24 h).

For antibody production, the stably transfected CHO cells
secreting the chimeric IgG and IgM, respectively, were cul-
tured in spinner flasks in CHO-S-SFM II medium, supple-
mented with HT supplement, until a cell density of about
1×10^6 cells/ml was reached. Following removal of the cells
from the cell culture supernatant by centrifugation
(400 × g, 15 min), the chimeric antibody was purified using
a protein A column (HiTrap r-protein A FF, Amersham Phar-
macia Biotech) for chimeric IgG or an anti-human Fc5μ anti-
body affinity column. The purified antibody fraction eluted
by sudden pH change was re-buffered in PBS and concentrated
using Centriprep centrifuge tubes (cut-off 50 kDa, Milli-
pore).

6. Sequence adaptation of the core 1-specific antibody sequences to human germ line sequences

To adapt the core 1-binding antibody sequences to human se-
quences, a search for homologous sequences was conducted in
the data base of human germ line sequences, and humanized
core 1-binding sequences were developed using human consen-
sus sequences and findings concerning the canonical struc-
ture of human antibodies. The human germ line sequence
V_H1-46 was used as model for the variable heavy chain, and
the sequence of A18 for the variable light chain.

The humanized V_H and V_L sequences SEQ ID Nos. 56 to 79 and
85 to 94, respectively, were produced using a gene assembly
PCR (single-overlap extension PCR). The PCR reaction pro-
ceeded according to the following scheme: first denatura-
tion at 94°C for 2 min, followed by 30 cycles of denatura-
tion at 94°C for 45 s, annealing at 55°C for 45 s and elon-
gation at 73°C for 1.5 min, and finally, an elongation step
at 73°C for 7 min.

The V_H and V_L chains thus produced were cut using the enzymes NcoI and XhoI or NotI and XhoI and cloned into a cloning vector (pLitmus 28 and pBluescript KS, respectively) for sequencing. The proper V_H and V_L chains were subsequently re-amplified to insert a BbsI restriction site at the 3' end of V_H and at the 5' end of V_L in order to link V_H and V_L via the latter using only one alanine as linker. Following ligation, the complete scFv (the ligation products) were amplified using the flanking primers and cloned into a bacterial expression vector.

7. Specificity analysis of core 1-specific recognition molecules in an ELISA

Various carbohydrate-PAA conjugates (synthesomes) and glycoproteins were used as antigens: asialoglycophorin (AGP), glycophorin (GP) and asialofetuins (Sigma); PAA (poly[N-(2-hydroxyethyl)acrylamide] conjugates: Galβ1-3GalNAcα1-OC₃H₆NH-PAA and Galβ1-3GalNAcα1-p-OC₆H₄NH-PAA as core 1 (α-anomer) conjugates with varying linker lengths, Galβ1-3GalNAcβ1-OC₃H₆NH-PAA as β-anomer of core 1, Galα1-3GalNAcα1-OC₃H₆NH-PAA and Galα1-3GalNAcβ1-OC₃H₆NH-PAA as additional stereoanomers of core 1, the core 2 structure Galβ1-3(GlcNAcβ1-6)GalNAcα1-OC₃H₆NH-PAA and derivatives of GalNAcα1-OC₃H₆NH-PAA, Neu5Acα2-3Galβ1-3GalNAcα1-OC₃H₆NH-PAA, Galβ1-3(Neu5Acα2-6)-GalNAcα1-OC₃H₆NH-PAA, GlcNAcβ1-2Galβ1-3GalNAcα1-OC₃H₆NH-PAA, GlcNAcα1-3Galβ1-3GalNAcα1-OC₃H₆NH-PAA, GalNAcα1-3Galβ1-OC₃H₆NH-PAA and 3'-O-Su-Galβ1-3GalNAcα1-OC₃H₆NH-PAA.

Using the respective stock solutions (1 mg in 1 ml of bidist. H_2O) stored in portions at -20°C, a dilution of 5 µg/ml in PBS was produced. 50 µl/well of the above was pipetted in a microtiter plate (NUNCLON-TC Microwell 96 F), and the test plate was incubated at 37°C for 1 hour and at 4°C overnight. On the next day, the test plate was washed 3

times with PBS/0.2% Tween. Subsequently, non-specific bind-
ing sites were blocked with 2% BSA in PBS, and 50 µl of the
first antibody was applied (chimeric IgG and IgM, respec-
tively: 0.1 µg/ml, purified, in PBS/0.1% BSA or undiluted
culture supernatant of producing CHOdhfr- cells; multibod-
ies: 10 µg/ml in PBS/0.1% BSA). After three wash steps with
PBS/0.2% Tween, the corresponding secondary antibodies,
peroxidase-coupled, were employed (an anti-mouse or anti-
human Fcγ1 or µ antibody for complete antibodies, an anti-
His-tag antibody for multibodies) to detect the specifi-
cally bound antibody constructs. To detect the bound secon-
dary antibody, a color reaction with TMB (3,3',5,5'-
tetramethylbenzidine) was performed. After 15 minutes the
reaction was quenched by adding 2.5 N H_2SO_4. Measurement was
performed using a microtiter plate photometer with 450 nm
filter in dual mode versus 630 nm reference filter.

Representative results are illustrated in Figures 5 and 6.
In Figure 5, two recognition molecules with varying loop
sequences in IgM format are compared. The antibody con-
structs mIgM-Karo2 (SEQ ID NO. 107 and SEQ ID NO. 109) and
mIgM-Karo4 (SEQ ID NO. 108 and SEQ ID NO. 110) bind to the
core 1 antigen in a highly specific fashion, preferably to
the α-anomer, Galβ1-3GalNAcα, and more weakly to the
β-anomer, Galβ1-3GalNAcβ. It is also possible that the rec-
ognition molecules of the invention bind the α-anomer Galβ
1-3GalNAcα only, or both anomers Galβ1-3GalNAcα and
Galβ1-3GalNAcβ in the same way. In addition, mIgM-Karo4
binds the core 2 structure Galβ1-3(GlcNAcβ1-6)GalNAcα. None
of the other tested carbohydrate structures, not even
structurally closely related structures, are recognized by
the binding proteins claimed herein. Being a core 1-bearing
glycoprotein, AGP shows a strong signal with both variants,
and the asialofetuin glycoprotein - likewise bearing core 1
- reacts significantly stronger with the Karo2 variant,
this very likely being related to the different core 1 den-

sity in the two proteins. Figure 6 shows the specificity pattern of the humanized recognition molecules, selected in an exemplary fashion, Karo11 (SEQ ID NO. 56 and SEQ ID NO. 90), Karo21 (SEQ ID NO. 59 and SEQ ID NO. 90) and Karo38 (SEQ ID NO. 69 and SEQ ID NO. 90) with varying framework sequences in scFv format and with one amino acid as linker. In this case as well, the same specificity pattern is seen, as described in the definition of core 1-specific binding in the meaning of the invention (see above).

Specific binding of various preferred formats and combinations in ELISA, exemplified on AGP, GP and/or Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA, is illustrated in Figures 7a through e.

8. Immunohistologic and immunocytologic staining

For immunohistologic staining, frozen sections of appropriate tissue samples were air-dried and fixed with 10% formaldehyde in PBS for 15 min. To reduce the endogenic peroxidase activity, the sections were treated with 3% hydrogen peroxide in PBS and, following blocking of non-specific binding sites with pre-absorbed rabbit serum on neuraminidase-treated erythrocytes, incubated with a core 1-specific primary antibody. Subsequently, the preparations were incubated with an appropriate secondary antibody (anti-mouse or anti-human IgG or IgM, POD-coupled). The staining reaction was performed using the peroxidase substrate diaminobenzidine, and counter-staining with hematoxylin.

The exemplary recognition molecule mIgM-Karo4 according to the invention undergoes reaction with only a very small number of structures in normal tissue. However, said structures are located in areas inaccessible to an antibody (Table 3).

Table 3:

Reaction of human normal tissue with the core 1-specific
mIgM-Karo4 antibody

Type of tissue	Reactivity
Epidermis - basal membrane	negative
Stomach	
Foveola epithelium	negative
Fundic glands	negative
Corpus glands	negative
Colon mucosa	negative
Spleen	
Splenic trabeculae	negative
Reticular cells	negative
Lymphocytes	negative
Endothelium	negative
Prostate	negative
Liver	
Hepatocytes	negative
Kupffer cells	negative
Bile tract	negative
Lymphatic nodes	
Lymphocytes	negative
Reticular cells	negative
Gall bladder	negative
Adrenal gland	
Adrenal cortex	negative
Adrenal medulla	negative
Bladder	negative
Heart	negative
Pancreas	
Glandular ducts	positive
Acini	negative
Islets of Langerhans	negative

The recognition molecules as claimed give positive reaction with a variety of carcinomas. The data in table 4 show that core 1-specific recognition molecules recognize a high percentage of tumor patients of one indication, which differs from one indication to the other.

Table 4:

Reaction of human tumor tissue with the core 1-specific mIgM-Karo4 antibody

Type of tissue	Reactivity
Colon carcinoma	
Primary carcinoma	31/52
Liver metastases	20/22
Lung carcinoma	
Large cell	3/8
Bronchoalveolar	1/1
Adenocarcinoma	6/6
Bladder carcinoma	5/9
Stomach carcinoma	
Intestinal type	8/8
Diffuse type	3/3
Prostate carcinoma	9/9
Mammary carcinoma	
Intraductal/ductal	8/10
Slightly differentiated	2/5
Mucinous	1/1
Thyroid carcinoma	0/10
Adrenal carcinoma	
Clear cell	4/9
Transitional cell	2/5
Cervical carcinoma	1/2
Ovarian carcinoma	
Adenocarcinoma	2/2
Endometrioid	2/2
Teratoma	2/2
Glioblastoma	0/3

To develop a mouse tumor model, various xenotransplants were investigated. The xenotransplants were human colon

carcinoma tissues repeatedly passaged on nude mice. In an exemplary fashion, Figure 8 shows immunohistochemic staining of a xenotransplant preparation with the core 1-specific cIgG-Karo4 antibody.

Immunofluorescence was used for the immunocytologic stainings. To this end, appropriate cells were slightly dried on microscope slides and fixed with 5% formaldehyde for 10 min. Following blocking of non-specific binding sites with BSA (1% in PBS), the cells were incubated with the primary antibody. This was followed by washing 3 times with PBS and incubation with the appropriate fluorescence-labelled secondary antibody (anti-mouse or anti-human IgG or IgM for complete antibodies; anti-myc-tag or anti-His-tag antibodies for single-chain antibody fragments). After repeated washing with PBS, the cells were embedded in Mowiol.

Various cell lines were tested with core 1-specific recognition molecules in immunofluorescence. A number of tumor cell lines, as well as some leukemia cell lines gave positive reaction (Table 5 and Fig. 9).

Table 5:
Reactivity of various cell lines with core 1-specific mIgM-Karo1 or mIgM-Karo4 antibodies

Cell lines	Reactivity
KG-1	positive
ZR-75-1	positive
T47D	(positive) few cells
U266	negative
LN78	positive
HT29	positive
HCT116	negative
HepG2	negative
K562	negative
NM-D4	positive

In an exemplary fashion, Figure 9 shows fluorescence labelling of KG-1 cells, an acute myeloid leukemia cell line, with various antibody constructs, a murine IgM, and two scFv antibodies with different linker length (SEQ ID NO. 95 with 18 amino acids and SEQ ID NO. 104 with one amino acid as linker). All three constructs show specific staining of the tumor cell line, the monovalent antibody fragment SEQ ID NO. 95 showing the weakest signal.

9. Chelating and radioactive labelling of antibodies and antibody fragments

Using conjugation, a chelating agent allowing binding of a radioactive metal was covalently bound to the cIgG-Karo4 antibody and to the multibody of sequence SEQ ID NO. 104, respectively. Commercial products from Macrocyclics (Dal-

las, USA), p-isothiocyanatobenzyl-diethylenetriaminepenta-
acetic acid (p-SCN-Bz-DTPA) and p-isothiocyanatobenzyl-
1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid
(p-SCN-Bz-DOTA) were employed as chelating agents. Both
chelating agents are suitable for linking to antibodies for
radiolabelling thereof [Brechbiel et al., 1986; Kozak et
al., 1989; Stimmel et al., 1995].

Conjugation proceeds via reaction of the isothiocyanate
group in the chelating agent with a free ε-amino group of
the amino acid lysine on the antibody, thus forming a cova-
lent N-C bond between chelating agent and antibody.

Initially, the purified antibody or the purified antibody
fragment must be re-buffered in coupling buffer, pH 8.7. To
this end, ultrafiltration in a filtration cartridge (Cen-
triprep of YM50 (Amicon)) was performed. This was done by
repeated dilution with a 10fold volume and filtration
through a membrane of defined pore size using centrifuga-
tion. In this way, PBS was replaced by alkaline coupling
buffer (0.05 M sodium carbonate, 0.15 M sodium chloride, pH
8.7).

Chelating was performed using the bifunctional chelating
agents p-SCN-Bz-DTPA and p-SCN-Bz-DOTA, respectively. For
the chelating reaction, the protein (1 to 10 mg/ml) in cou-
pling buffer and a solution of chelating agent of 1 mg/ml
in 2% DMSO/water were mixed such that a molar excess of
chelating agent was ensured. This was followed by incuba-
tion of the mixture at 37°C for 1 hour. Subsequently, non-
bound chelating agent was removed by ultrafiltration in the
same vessel (Centriprep YM50 (Amicon)) and, as described
above, this was re-buffered to pH 4.2 in a loading buffer
(0.15 M sodium acetate, 0.15 M sodium chloride, pH 4.2) re-
quired for radioactive labelling. The protein concentration

during and after this step was re-adjusted to 1-10 mg/ml using UV measurement at 280 nm.

Conditions for the chelating reaction had to be found, which would allow radiolabelling of the antibody without substantially reducing the bioactivity thereof.

The chelated antibody was loaded with a radioactive metal, thereby producing the radioantibody. The isotopes [111]indium and [90]yttrium were used for loading. Both have comparable chemical and physicochemical properties, being bound as trivalent ions ([111]In^{3+}, [90]Y^{3+}) by the chelating agent. The antibody labelled with [111]indium is a γ-emitter and is used clinically to find the individual dose for a patient, while [90]yttrium is a β-emitter which is used therapeutically. The half-lives are 67 hours for ^{111}In and 64 hours for [90]Y.

[111]Indium chloride from the company NEN (Perkin Elmer, Belgium) was used for loading. The radioactive metal is supplied in a solution of hydrochloric acid. First of all, the [111]InCl$_3$ solution was brought to an HCl concentration of 1 M. Subsequently, this was diluted with 0.05 M HCl to a specific activity of 80-320 mCi/ml, and an aliquot thereof was used for incorporation in the chelated antibody, in which case the added volume of HCl-acidic [111]InCl$_3$ solution should be equal to the volume of antibody solution supplied in the coupling buffer of pH 4.2 so as to ensure pH stability. The incubation time was 1 hour at 37°C, with occasional careful mixing.

Subsequently, the filter insert was re-inserted into the filtration cartridge and re-buffered as described above in phosphate buffer, pH 7.2, including a physiological content of sodium chloride, thereby effecting separation of high-molecular weight radiolabelled antibody and unbound [111]InCl$_3$. Quantification of ^{111}In incorporation in the chelated anti-

body was performed using thin layer chromatography. The incorporation rate of radioactive metal was 70-99% of the radioactivity employed.

10. Detection of core 1-positive, secretory MUC1 in a sandwich ELISA

Core 1-positive, secretory MUC1 can be detected in a sandwich ELISA. A MUC1-specific antibody was used as scavenger antibody of MUC1, and a core 1-specific antibody to detect the core 1 antigen. A third enzyme- or fluorescence-coupled antibody must be used to detect the secondary antibody.

The supernatants of two tumor cell lines (K562 and T47D) were analyzed as examples. The results are illustrated in Table 6. 10^5 cells per ml of cell culture medium were seeded, cultured for 4 days without replacing the medium, an aliquot was subsequently drawn, and the cell culture supernatant was separated from the cell pellet by centrifugation. 50 µl of undiluted supernatants were used in the ELISA. The anti-MUC1-anti-core 1 sandwich ELISA was carried out by coating the microtiter plate with scavenger antibody (1 µg/ml) in PBS at 4°C overnight. Three different concentrations of antibody were used for coating (1 µg/ml, 2 µg/ml and 4 µg/ml). The 1 µg/ml coating was found to be the most sensitive in the sandwich ELISA. Subsequently, the coated plates were washed twice with PBS and blocked in 5% BSA, 0.05% Tween 20 in PBS for 1.5 hours at room temperature. The blocking buffer was removed, the plates were washed once more with 0.1% Tween 20 in PBS (washing buffer), the samples were added and incubated at room temperature for 1.5 hours. Cell culture medium or 2% BSA in washing buffer (dilution buffer for secondary antibody) was used as negative control. Positive control was not available. After washing three times, neuraminidase treatment was performed in the wells intended for that purpose. To

this end, a neuraminidase solution (DADE Behring, Germany) was diluted 1:5 in imidazole buffer (0.68 g of imidazole, 0.19 g of $CaCl_2$ and 0.4 g of NaCl in 100 ml of H_2O, pH 6.8) and incubated at 50 µl/well for 30 min at 37°C. As a control, the imidazole buffer with no neuraminidase solution was incubated in a corresponding well. Subsequently, the wells were washed three times, and the mIgM-Karo4 antibody for the detection of core 1 antigen was added at a dilution of 1:500 in 2% BSA in washing buffer and incubated at room temperature for another hour. Again, this was washed three times, followed by addition of a peroxidase-coupled anti-mouse IgM(µ) antibody (Dianova) diluted 1:5000 in 2% BSA washing buffer and incubation for 1 hour at room temperature. Finally, the plates were washed twice in washing buffer and once in PBS. The staining reaction was performed in 25 mM citric acid, phosphate buffer, pH 5.0, with 0.04% H_2O_2 and 0.4 mg/ml o-phenylenediamine (Sigma) in the dark at room temperature. The staining reaction was quenched by adding 2.5 N sulfuric acid (final concentration 0.07 N) and measured in an ELISA Reader at 492 nm with a 620 nm reference filter.

Table 6:
Analysis of core 1-positive MUC1 in culture supernatants of two cell lines with and with no neuraminidase treatment in a sandwich ELISA

Cell line	Signal	
	- NeuAcdase	+NeuAcdase
K562	–	+
T47D	+	+++

11. Effective binding of radiolabelled core 1-specific recognition molecules in tumor cells

The core 1-positive tumor cell line NM-D4 [DSMZ deposit No. DSM ACC2605] (cf. Table 5) was used to test the binding capability of radiolabelled recognition molecules in core 1-positive tumor cells. In each double determination, a defined number of cells was placed in a 1.5 ml vessel and incubated with increasing amounts of antibodies. Following washing, the amount of bound antibodies was determined on the basis of the counting rate.

2×10^6 cells per batch are required. Following pre-incubation of the cells for one hour on ice, the required amount of cells was placed in reaction vessels, centrifuged (5 min at 1000 × g, 25°C), and the supernatant was removed. Thereafter, this was filled up with PBS/0.1% Tween20/1% BSA to make a volume of 200 µl, subtracting the amount of recognition molecules to be added later. Subsequently, the corresponding ^{111}In-labelled recognition molecule (see Example 9) was added to make a final volume of 200 µl (about 0.5 to 20 µg, depending on the recognition molecule), and the batch was incubated for one hour at 4-8°C. Following centrifugation (4 min, 1000 × g, 25°C), the supernatant was removed and the cell pellet carefully resuspended in 400 µl of PBST/1% BSA. After another wash, the cell pellet was measured in the vessel on a gamma counter. The specific counting rates were determined in the initial solutions of defined concentration, and the value in cpm/ng was used as a basis of relativizing the measured values of bound antibody. Free binding is obtained from the difference of total amount and amount of bound antibody. These values were plotted in a diagram as ratio of bound/non-bound versus bound amount, the slope in the linear region of the curve was determined, and the abscissa intersection was determined (Scatchard analysis). The abscissa intersection indi-

cates the number of binding sites/cell. The slope of the straight line furnishes the association constant K_{ass} in M^{-1}.

Figure 10 exemplifies the Scatchard analysis of binding of radiolabelled recognition molecules in scFv format with the sequence SEQ ID NO. 104 and with one amino acid as linker on NM-D4 cells (two different preparations).

Table 7 summarizes the association constants and the number of cell binding sites of different core 1-specific multibodies on NM-D4 cells.

Table 7:
Cell binding test and Scatchard analysis with ^{111}In-labelled recognition molecules on NM-D4 cells.

Antibody	K_{ass} $[M^{-1}]$	Number of binding sites/cell
SEQ ID NO. 105	1.1×10^7	4.8×10^6
SEQ ID NO. 104	2.1×10^6	8.1×10^6
SEQ ID NO. 103	1.2×10^6	9.2×10^6

12. Accumulation of radiolabelled core 1-specific recognition molecules on core 1-positive tumors in an *in vivo* tumor model

ZR-75-1 cells as tumor model were injected subcutaneously in nude mice (Ncr: nu/nu, female). After about 3-4 weeks, the tumor is palpable under the skin. To the tumor-bearing mice (n=4 per point in time) 5 µg of ^{111}In-labelled multibody (SEQ ID NO. 104 and SEQ ID NO. 105, respectively) in 200 µl was administered into the tail vein. After 24 hours the mice were sacrificed and the radioactivity distribution in the tumor, in serum and in organs was determined. Table 8 shows the specific high accumulation of multibodies in

the tumor (in %ID/g tumor, relative to injected dose and tumor weight) compared to serum and organs.

Table 8:

Biodistribution of ^{111}In-labelled recognition molecules in tumor-bearing mice

	SEQ ID NO. 104	SEQ ID NO. 105
Serum (%ID/ml)	1.4 ± 0.16	1.0 ± 0.24
Tumor (%ID/g)	10.8 ± 2.88	8.1 ± 1.45
Liver (%ID/g)	3.7 ± 0.15	5.3 ± 0.92
Lung (%ID/g)	1.7 ± 0.11	1.9 ± 0.19
Heart (%ID/g)	1.5 ± 0.06	1.9 ± 0.19
Spleen (%ID/g)	5.4 ± 0.75	6.7 ± 1.07
Brain (%ID/g)	0.1 ± 0.01	0.1 ± 0.00
Bone marrow (%ID/g)	1.0 ± 0.16	1.7 ± 0.90

13. Therapeutical study for reduction of core 1-positive tumors with radiolabelled core 1-specific recognition molecules in an *in vivo* tumor model

The therapeutical studies were carried out using the same established ZR-75-1 tumor model as described in the biodistribution studies (see Example 12). To this end, the chelated recognition molecules (see Example 9) were loaded (pH 4.5, 37°C, 30 min; cf. [111]indium incorporation) with [90]yttrium (a β-emitter to destroy the tumor cells), and the stability was controlled using thin layer chromatography. The tumor-bearing mice (about three weeks after subcutaneous injection of ZR-75-1 cells) were given 200 µl into the tail vein. The injection solution included the ^{90}Y-labelled multibody (up to a maximum of 100 µCi per dose) in Ca/Mg-PBS with 0.2 to 4% fetal calf serum to protect against radioly-

sis. Control groups received the same injection with no radioactively labelled recognition molecule. Body weight and tumor size were measured twice a week and compared. The relative tumor growth was determined considering the respective tumor size at the beginning of treatment. A second injection was given three weeks after the first treatment. Significant reduction in tumor growth compared to the control group was possible by suitable treatment.

Figure legends

Fig. 1a: **Sequences of linkers in various multibody single-chain antibody fragments.**

Fig. 1b: **Cloning diagram for the preparation of single-chain antibody fragments having different linker length.**

Fig. 2:   **Vector for cloning and bacterial expression of single-chain antibody fragments.**

Fig. 3:   **Analysis of multibodies in scFv format with varying linker length in ELISA.**

Multibodies having the amino acid sequences SEQ ID Nos. 95, 96, 97, 98, 99, 100, 101, 103, 104 and 105 were expressed in *E. coli* as described above and the periplasm fractions obtained. Asialoglycophorin, which is a core 1-bearing glycoprotein, was used as antigen in the ELISA. Step-by-step linker length reduction results in increased binding to asialoglycophorin. The best binding properties are seen in the variants having SEQ ID Nos. 104 and 105. These multivalent constructs in dia/triabody format are preferred embodiments of the invention.

Fig. 4:   **Vector system for cloning and eukaryotic expression of chimeric antibodies in IgG1 or IgM format.**

Fig. 5, 6: Specificity analysis in ELISA.

Various glycoproteins and carbohydrate-PAA conjugates were used as antigens. Asialoglycophorin **[1]**; glycophorin **[2]**; asialofetuins **[3]**; Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[4]**; Galβ1-3GalNAcα1-p-OC$_6$H$_4$NH-PAA [5]; Galα1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[6]**; Galβ1-3GalNAcβ1-OC$_3$H$_6$NH-PAA **[7]**; Galα1-3GalNAcβ1-OC$_3$H$_6$NH-PAA **[8]**; Galβ1-3(GlcNAcβ1-6)GalNAcα1-OC$_3$H$_6$NH-PAA **[9]**; GalNAcα1-OC$_3$H$_6$NH-PAA **[10]**; Neu5Acα2-3Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[11]**; Galβ1-3(Neu5Acα2-6)GalNAcα1-OC$_3$H$_6$NH-PAA **[12]**; GlcNAcβ1-2Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[13]**; GlcNAcα1-3Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[14]**; GalNAcα1-3Galβ1-OC$_3$H$_6$NH-PAA **[15]**; and 3'-O-Su-Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA **[16]**. BSA [17] was used as control. In **Figure 5**, two antibodies in IgM format with varying CDR sequence composition were used. **Figure 6** shows the specificity pattern of three humanized recognition molecules in scFv format with varying framework sequences.

Fig. 7:   **Specific binding of different preferred formats and combinations of recognition molecules of the invention in ELISA, with AGP, GP and/or core 1-PAA (Galβ1-3GalNAcα1-OC$_3$H$_6$NH-PAA) antigens as examples.**

Fig. 8:   **Immunohistochemical staining of xenotransplant preparations.**

Human colon carcinoma tissue was transplanted on nude mice and passaged after reaching a specific size. The tumor tissue was embedded and dissected and used in immunohistochemical staining. In   a) the tissue was labelled with

cIgG-Karo4 as primary antibody and an anti-human Fcγ anti-body, POD-coupled, as secondary antibody. Brown staining characterizes core 1-positive structures.

Fig. 9:   **Fluorescence-labelling of cells of the KG-1 tu-mor cell line with different core 1-specific recognition molecules.**

Fig. 10:   **Scatchard diagram for analysis of cell binding of radiolabelled core 1-specific recognition molecules.** Binding data of multibody SEQ ID NO. 104 with a linker length of one amino acid are illustrated in an exemplary fashion (Pr1 and Pr2 correspond to two different preparations). B: amount bound to cells [M]; F: free binding as difference of total and bound amount of an-tibody [M]. The corresponding straight-line equation is given at the top, the slope of the straight-line representing the association con-stant.